Policy Forum

Association between Macroscopic-factors and Identified HIV/AIDS Cases among Injecting Drug Users: An Analysis Using Geographically Weighted Regression Model*

XING Jian Nan¹,², GUO Wei¹,³, QIAN Sha Sha¹, DING Zheng Wei¹, CHEN Fang Fang¹, PENG Zhi Hang², QIN Qian Qian¹, and WANG Lu¹,⁹

Drug use (DU), particularly injecting drug use (IDU) has been the main route of transmission and spread of Human Immunodeficiency Virus (HIV)/Acquired Immune Deficiency Syndrome (AIDS) among injecting drug users (IDUs)[1]. Previous studies have proven that needles or cottons sharing during drug injection were major risk factors for HIV/AIDS transmission at the personal level[2-4]. Being a social behavioral issue, HIV/AIDS related risk factors should be far beyond the personal level. Therefore, studies on HIV/AIDS related risk factors should focus not only on the individual factors, but also on the association between HIV/AIDS cases and macroscopic-factors, such as economic status, transportation, health care services, etc[1]. The impact of the macroscopic-factors on HIV/AIDS status might be either positive or negative, which are potentially reflected in promoting, delaying or detecting HIV/AIDS epidemics.

China has reported HIV/AIDS infection among IDUs since domestic HIV/AIDS cases were first identified among them in Yunnan Province in 1989[5]. In response to HIV/AIDS epidemic, China has established a real name based HIV/AIDS case reporting system (CRS) since 1985[6]. All hospitals and centers for diseases control and prevention (CDCs) are obligated to report identified HIV/AIDS cases. According to the legislation, once a HIV/AIDS case was identified by Western Blot assay, an interview will be conducted by health care personnel to provide counseling to him/her within 10 work days. A case report form (CRF) should be completed and delivered to the National Center for AIDS/STD Control and Prevention (NCAIDS) after a face to face interview. Demographic information (age, gender, occupation, address, etc.) and the history of high risk behaviors of the interviewed cases were collected based on CRF. And the most likely route of HIV infections was judged by healthcare personnel according to the history of their high risk behaviors. By collecting demographic information of people living with HIV/AIDS, CRS provided both the number of HIV/AIDS cases and the time and their spatial distribution, which enabled us to explore the spatial distribution characteristics of HIV/AIDS cases among IDUs[7].

During spatial analysis, autocorrelation, instability and heterogeneity of the spatial data should be considered carefully[8-9]. General spatial autocorrelation is a technique which is used to detect the spatial cluster of data, measure and analyze the degree of dependency among observations in a geographic space[8]. In order to analyze the variance and instability of spatial parameters, an effective and sample technique-geographically weighted regression (GWR) was initiated by Brunsdon et al., which allowed the value of coefficients to change at different spatial sites. GWR is produced to generate parameters disaggregated by the spatial units of analysis by local spatial regression. This allows assessment of the spatial heterogeneity and instability in the estimated associations between the independent and dependent variables by expanding ordinarily linearity regression through embedding spatial data structure into the regression model, and in this way the GWR model can reflect the spatial instability of the coefficients at different spaces[9-12].

In this study, we focus on the spatial autocorrelation of HIV/AIDS cases among IDUs in China from 2007 to 2011, and explore the association between the spatial distribution and

doi: 10.3967/bes2014.055

1. National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, China; 2. Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 210029, Jiangsu, China
macroscopic socio-economic factors, instead of personal level factors, by using general spatial autocorrelation and GWR model.

**Data Management**

HIV/AIDS positive IDUs, identified by CRF from 2007 to 2011, were downloaded from the web based China's National HIV/AIDS CRS, which was managed by the NCAIDS, China CDC. Name linked information, such as name, ID number, home address, etc., were deleted before data analysis. National standard geocode at the provincial level was used to identify which province the HIV/AIDS positive IDUs came from.

Population size and twenty-one original demographic and socio-economic indicators, including economic status, transportation, social security and healthcare services, etc., were collected at the provincial level from China Statistical Yearbook from 2007 to 2011, released by the National Bureau of Statistics of China. To demonstrate the average level of these socio-economic indicators during 2007-2011, arithmetical mean of these indicators during the five years was used to extract four synthesized variables by principal component analysis.

**Variable Definition**

Identified HIV/AIDS cases density for each province was the key variable in spatial autocorrelation analysis and acted as the dependent variable in GWR. Identified HIV/AIDS cases density was calculated at the provincial level as below to reflect the HIV/AIDS epidemic situation among IDUs.

\[
\text{Identified cases density} = \frac{\text{Identified HIV/AIDS cases among IDUs from 2007-2011}}{\text{mean of population size from 2007-2011}/10^6}
\]  

By using principal component analysis, four synthesized variables were extracted from 21 original socio-economic indicators to reflect economic status, transportation, social security and health care services in each province. These four provincial synthesized variables acted as independent variables in GWR.

**Principal Component Analysis**

Principal component analysis using the method of dimension reduction, not only can extract a principal component from many variables in the premise of losing little information of variables, but can also handle the possible multicollinearity among variables. Cumulative variance contribution is a statistic parameter which assesses the result of principal component analysis, when the cumulative variance contribution≥70%, which denotes that the result is reliable and the extracted principal component can represent most of the information of variables. Factor loading represents the association between the extracted principal component and variables; the higher absolute value of factor loading is, the more information about variables is represented by the extracted principal component.

In this paper, we used principal component analysis to extract four provincial synthesized variables (macroscopic-factors), from each four genres of original variables (indicators): economic status, means of transportation, social security, and health care services. These four synthesized variables were expected to represent the overall situation and information which were interesting to us. The process of principal component analysis was conducted by SPSS20.0 and the four macroscopic-factors were embedded in the GWR model as independent variables.

**General Spatial Autocorrelation**

We used the spatial code in the record of each HIV/AIDS cases among IDUs to match the geographic information system. It was assumed that the provinces in the whole country did not differ from each other, and autocorrelation was applied to analyze the country as a whole in this paper. Moran’s Index was used to analyze the general spatial autocorrelation and the value of Moran’s Index was between [-1, 1]. When the value of Moran’s Index>0 and Z-value>1.96 or the Moran’s Index <0 and Z-value <-1.96, it represented clustered HIV/AIDS cases; if the value of Moran’s Index was close to zero and Z-value was between -1.96 and 1.96, then it represented randomly distributed HIV/AIDS cases [6]. In this study, we conducted general spatial autocorrelation for identified cases density. If P<0.05 and Moran’s Index≠0, it indicated that HIV/AIDS cases were clustered among IDUs in China, but the data were not independent, and it was too serious to use the traditional statistical model to analyze them, and so the GWR model was better to be used for analyzing these data [9].

**Geographically Weighted Regression Model**

The GWR model allows the value of coefficients to change at different spatial sites (in this paper, spatial sites were defined as 31 provinces) and it is
capable to show the local association between the dependent variable and the independent variable, which mean that the GWR model can detect the instability of spatial data. The Ordinarily Linearity Regression Model is:

$$y_i = \beta_0 + \sum_k \beta_k X_{ik} + \epsilon_i$$

but the GWR model expands it to:

$$y_i = \beta_0 (u_i, v_i) + \sum_k \beta_k (u_i, v_i) X_{ik} (u_i, v_i) + \epsilon_i$$

$$(u_i, v_i)$$ is the center-of-mass-coordinate points of spatial site $$i$$, which is used as geographically weighted. $$\beta_k (u, v)$$ is the value of continuous coefficient $$\beta_k$$ at the spatial site $$j^{[11-12]}$$. In this study, the dependent variable was always identified as infection density at the provincial level, while the independent variables were four macroscopic-factors that were extracted from principal component analysis. The GWR model can indicate spatial variance, namely spatial instability, of the association between the provincial identified cases density among IDUs and four macroscopic-factors through a continuous parametric surface. That means we could obtain different values of coefficients of four macroscopic-factors in 31 provinces. We used ArcInfo software (ESRI® ArcMap™, Version 10.0) to operate the GWR model. In the output of this model, Condition Number ($$\text{Cond}$$) indicates local collinearity between macroscopic-factors; when all of them <30, the result of the GWR model is reliable and accurate.

After stimulation by using the GWR model, a general spatial autocorrelation test for residuals was conducted and if there was no general spatial autocorrelation observed in the whole area, the fitness of this model was good; otherwise, the fitting effect was unsatisfactory.$^9$

**Basic Information**

The identified cases density among IDUs in China from 2007 to 2011 was 4.49/100 000 persons. Provincial identified cases density was significantly different in 31 provinces, from 0.083/100 000 persons in Jilin Province to 35.979/100 000 persons in Xinjiang Uygur Autonomous Region. Figure 1 shows a map with the distribution of identified cases density among IDUs in 31 provinces during this period. Although all of the 31 provinces reported infection cases among IDUs in these years, Xinjiang Uygur Autonomous Region, Guangxi Zhuang Autonomous Region, Yunnan and Sichuan provinces and were identified as highly endemic.

**Exaction of Macroscopic-factors**

We extracted synthesized variables (macroscopic-factors) from original indicators of the economic status, transportation, social security and health care services. The variance contribution of the economy-
factor was 76.60%, the transportation-factor was 64.98%, the social security-factor was 85.06% and the health care service-factor was 70.47%, which meant that four extracted factors represented most of the information contained in their own original indicators. In addition, the absolute values of factor loading were relatively large, which also indicated that information about original indicators was represented by macroscopic-factors well. Provincial original indicators and results of principal components analysis are shown in Table 1.

**Spatial Autocorrelation**

Using a distance matrix to conduct general spatial autocorrelation for provincial identified cases density among IDUs, and, as shown in Table 2, the value of Moran’s Index was 0.102, while Z-value=2.472 and P-value=0.0134. Further general spatial autocorrelation test was conducted in respective years (2007-2011), and the results showed that provincial identified cases density among IDUs was clustered in the whole area for each individual year from 2007 to 2011. Although there was a decreasing trend in the value of Moran’s Index, indicating a reduction in the clustering attribute of HIV/AIDS cases among IDUs in China, it in no way suggested that traditional statistical regression model was suitable for analyzing these data.

<table>
<thead>
<tr>
<th>Original Socio-economic Indicators (unit)</th>
<th>Factor Loading</th>
<th>Macroscopic-factors</th>
<th>Variance Contribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Average remuneration of labor employment in urban areas (Yuan)</td>
<td>0.429</td>
<td>Economy</td>
<td>76.60</td>
</tr>
<tr>
<td>2. Consumption level of residents (Yuan)</td>
<td>0.497</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Urban residents' disposable income (Yuan)</td>
<td>0.502</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. GDP (Hundred million Yuan)</td>
<td>0.263</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Per capita net income of rural households (Yuan)</td>
<td>0.496</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Passenger transport volume (Ten thousand Persons)</td>
<td>0.491</td>
<td>Transportation</td>
<td>64.98</td>
</tr>
<tr>
<td>7. Passenger turnover volume (Million kilometers*Persons)</td>
<td>0.527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Civilian number of cars (Ten thousand)</td>
<td>0.480</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Length of Railways (Kilometer)</td>
<td>0.230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Grade Road length (Kilometer)</td>
<td>0.443</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. The number of urban workers participating in the basic medical insurance (Ten thousand Persons)</td>
<td>0.431</td>
<td>Social security</td>
<td>85.06</td>
</tr>
<tr>
<td>12. The number of rural residents participating in social endowment insurance (Ten thousand Persons)</td>
<td>0.419</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. The number of person participating in unemployment insurance (Ten thousand Persons)</td>
<td>0.429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Financial investment in social security and employment (Hundred million Yuan)</td>
<td>0.336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Financial investment for public safety (Hundred million Yuan)</td>
<td>0.428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Financial investment for national defense (Hundred million Yuan)</td>
<td>0.395</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Number of health institutions (Count)</td>
<td>0.473</td>
<td>Health care services</td>
<td>70.47</td>
</tr>
<tr>
<td>18. Number of beds in medical institutions (Count)</td>
<td>0.520</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Number of health workers per thousand people (Person)</td>
<td>-0.079</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Hospital visits (Ten thousand Persons)</td>
<td>0.492</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Financial investment in health care services (Ten thousand Yuan)</td>
<td>0.506</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* △represents unavailable number of health workers per thousand people in 2007.
**Geographically Weighted Regression Model**

Four provincial macroscopic factors were introduced into the GWR model as the independent variables and provincial identified cases density among IDUs in China from 2007 to 2011 acted as the dependent variable. The condition numbers of 31 provinces were all less than 30, indicating no local collinearity among four macroscopic factors. R-square of GWR model was 77.1%, and adjusted R-square was 62.5%. General spatial autocorrelation test for residual showed that the value of *Moran’s I* was 0.030 (Z=1.970, P=0.056), which indicated that the fitting effect of the GWR model was satisfactory. The economic status had a negative effect on the identified cases density among IDUs, and the coefficient for each province increased from -9.7 in western provinces to 0.1 in northeastern provinces. Transportation was another negative factor for the identified cases density among IDU for most of the provinces, except Xinjiang, with the coefficient value varying from -15.27 to 20.73. Social security was a positive factor for HIV/AIDS prevalence in all provinces, with the coefficients varying from 0.05 to 16.34. In addition to this, the primary health infrastructure also had the positive impact on identifying HIV positive cases among IDUs, and the coefficients generally decreased from east to west. The value of provincial coefficients of four macroscopic factors in the GWR model were shown in Figure 2.

<table>
<thead>
<tr>
<th>Year</th>
<th>Moran’s Index</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>0.194</td>
<td>3.318</td>
<td>0.0009</td>
</tr>
<tr>
<td>2008</td>
<td>0.161</td>
<td>2.875</td>
<td>0.0040</td>
</tr>
<tr>
<td>2009</td>
<td>0.159</td>
<td>2.832</td>
<td>0.0046</td>
</tr>
<tr>
<td>2010</td>
<td>0.155</td>
<td>2.781</td>
<td>0.0054</td>
</tr>
<tr>
<td>2011</td>
<td>0.142</td>
<td>2.605</td>
<td>0.0091</td>
</tr>
<tr>
<td>Total</td>
<td>0.102</td>
<td>2.472</td>
<td>0.0134</td>
</tr>
</tbody>
</table>

Table 2. Results of General Spatial Autocorrelation from 2007 to 2011

![Economy-factor (-9.7~0.1)](image1)

![Transportation-factor (-15.27~20.73)](image2)

![Social Security-factor (0.05~16.34)](image3)

![Health Care Services-factor (-27.15~8.47)](image4)

Figure 2. Coefficients of macroscopic factors in the GWR model.
Since 146 HIV/AIDS cases among IDUs were identified at border areas in 1989[14], DUs and IDUs have become the most HIV/AIDS affected populations in China. The researchers have proven the risk behaviors for HIV infection among IDUs[2-4], but few studies were conducted to identify the association between cluster characteristics of the HIV/AIDS epidemic among IDUs and the social macroscopic factors, such as economy, transportation, social security conditions and healthcare services. In this study, four social macroscopic-factors were extracted from dozens of socio-economic indicators to reflect the social environment at the provincial level by using principal component analysis. And then further geographical weighted model was built to identify the association of the numbers of reported HIV/AIDS cases among IDUs with these four macroscopic-factors.

The distribution of HIV/AIDS positive IDUs in China was reported earlier in several researches, and the clustering characteristics were related to the heroin trafficking route from southwestern and western provinces to the inland provinces of this country[15-16]. The spatial attributes of the HIV/AIDS epidemic among IDUs required the using of the geographical weighted model during impact factor analysis. In our study, we used the GWR model to analyze the potential association between the HIV/AIDS case density among IDUs and four genres of the original indicators that contained spatial attributes, including economy, transportation, social security, health care services.

We screened original indicators according to related literatures in the hope that the chosen variable could cover the real conditions of the urban and rural area from four genres[18-20]. Finally, 21 original indicators were chosen for our study. In order to reduce the number of variables and collinearity between variables, principal components analysis was used to extract four macroscopic-factors from variables. The value of variance contribution and factor loading indicated that the four macroscopic-factors could represent most information of variables selected earlier.

After fitting GWR model, the coefficients of economy-factor represented an increasing trend from the west to the east, the coefficients were in the range of -9.7 to 0.1. The coefficients of economy-factor were negative except for Heilongjiang and Jilin provinces, which indicated that with the development of economy, the identified cases density among IDUs would reduce in China.

The HIV/AIDS infection tends to begin and spread in impoverished countries, for example, Zambia, South Africa and other countries in Africa[21-22]. The spread of HIV/AIDS ruined most of economic progresses that these countries ever made. In our study, some western provinces in China, for example, Tibet and Xinjiang's absolute values of the coefficients of the economy were bigger than those of the inland provinces. There are two possible explanations for this phenomenon. First, eastern provinces have better economic development than western provinces, which means that improvement of the economy might produce more extensive, positive and meaningful influence on western provinces and it might reduce the infection density among IDUs. Second, because these provinces share the border with 'Golden Triangle', the number of infection cases is much more than in other provinces and despite similar level of economic development, which might also bring more positive influence on the western provinces. As for the positive coefficient of Heilongjiang (0.09) and Jilin (0.02), it might be due to lower level of their economic development than that of their neighboring provinces (Liaoning) and the identified cases density among IDUs was lower too. The improvement of economic level might restrain the spread of HIV/AIDS among IDUs in China to some degrees while the spread of HIV/AIDS might have the negative impact on economy in turn, and this impact can be observed in both direct and indirect ways. The direct impact can be the fiscal burden and medical expenses for society and the indirect impact can be loss of labor force[23].

Coefficients of transportation were negative in 30 provinces, but the coefficient of Xinjiang Uygur Autonomous Region was positive. The values of coefficients were incremental from the western provinces to the eastern provinces, and the value of the coefficient was in the range from -15.27 to 20.73. This result was different from one study published two years ago[3]. However, we integrated five original indicators and used principal components analysis to extract the macroscopic-factor of transportation, rather than one specific variable (Length of railways), indicating that our study contained more information regarding the condition of transportation in China. Further, just focusing on the length of railways will bring some biases into the study, because it ignored the other kinds of transportation means in China, for example, traveling by cars, which might be the dominant vehicle of transportation in China. In this paper,
although the convenience of traffic can lead to the increase of floating population and makes drugs circulation more convenient, under the control of the economy-factor and other macroscopic-factors, transportation-factor was still negative in 30 provinces, indicating that poor transportation conditions might make IDUs more prone to infection. Especially in Yunnan, Sichuan and other southwestern provinces, the development of transportation might significantly reduce the number of HIV/AIDS cases among IDUs. However, Xinjiang is located in the hub of drug trade route which begins from eastern Burma to northwestern China and then crosses the Chinese border to Kazakhstan\(^{[15]}\). The farthest distance for transporting drug requires good traffic. Logically, the better condition of transportation might lubricate the circulation of drugs, leading to infection among IDUs. Besides, given the vastness of Xinjiang Uygur Autonomous Region, its local transportation is behind most of provinces in China, for example, the mean of annual passenger transport volume among five years was only 314 780 thousand persons , whereas the highest volume was 4 120 180 thousand persons in Guangdong province, which also indicated that better transportation might create much more opportunities for transporting drugs and help to explain why the impact of transportation in Xinjiang was opposite to other provinces.

The coefficients of social security were positive in 31 provinces, which might indicate the higher level of social security and safety, the more reported HIV/AIDS cases among IDUs. There are two explanations for this phenomenon: (1) Because of the increasing number of urban or rural workers participating in the basic medical insurance and more financial investment for social security and employment, more and more IDUs have the opportunity to have access to HIV/AIDS antibody test, and as a result, more IDUs who were affected with HIV/AIDS were reported and identified. (2) The clustering of HIV/AIDS cases among IDUs spurs local government to make larger investment and formulate promising policy to prevent IDUs and other venerable populations from HIV/AIDS. However, the number of reported HIV/AIDS cases among IDUs and the positive rate of HIV/AIDS antibody in sentinel points of some traditional HIV/AIDS epidemic provinces both represented the decreasing trend\(^{[24-26]}\), which indicated that the first reason is more reliable and more logistical support for this study. Just like the coefficient of social security, the coefficients of health care service-factor were also positive in the whole country, except for Xinjiang and Tibet. The reasons behind this phenomenon might be similar to the social security-factor, due to an increasing number of medical and health institutions which might enable IDUs to gain better access to HIV/AIDS antibody test. Xinjiang and Tibet regions’ coefficient of health care services were -27.14 and -3.66, owing to the fact that health care services in Xinjiang and Tibet were much poorer than in other provinces. For instance, the number of health institutions in Tibet was only 6 602 in 2011. Given this situation, the better level of primary health infrastructure might reduce the risk for HIV/AIDS epidemic among IDUs in provinces where sanitation is poor, possibly owing to the improvement of medical environment and health education. Furthermore, the identified number of HIV/AIDS cases through injecting drug use was much higher in Xinjiang, leading to the lowest coefficient of health care services in the whole area.

The results of the present study have demonstrated that the IDUs infected with HIV and affected with AIDS are clustered in China, and the level of economic development and health care services are also not balanced at the provincial level, which makes the provinces have different responses to the association between macroscopic-factors and HIV/AIDS cases among IDUs. The number of infections among IDUs reported to the surveillance system was not only influenced by the real condition of infections but also by the strength of reporting and antibody testing. Correspondingly, macroscopic-factors influencing identified HIV/AIDS infection rate might be classified into two categories in our study: the internal factors and the external factors. From the results of the GWR model, the development of economy and smooth transportation might contribute to the reduction of HIV/AIDS prevalence among IDUs as an internal factor, while the higher level of social security and health care services might help local CDCs and hospitals detect and identify more infections as an external factor. However, we also should pay some attention to the values of coefficients of the macroscopic-factors. Because of using principal components analysis to extract four macroscopic-factors, values of the macroscopic-factors are much smaller than the values of the original indicators, and therefore HIV/AIDS prevalence among IDUs could not be easily reduced, and the epidemic could neither be
restrained in some key areas just by changing economy, transportation and other factors. Integrated interventions and programs at the individual level synchronized with improved macroscopic factors are strongly recommended[27].

There are some limitations in our study. First, we were not able to conduct the statistical test for the presence of spatial dependence at the individual level. However, the values of Cond, adjusted R-square and general spatial autocorrelation test for residual all indicated that the fitness of the GWR model was satisfactory. Second, our data have relied on national HIV/AIDS case reporting system and underreporting of HIV/AIDS cases may occur in some provinces for these or other reasons or HIV/AIDS cases among IDUs have not been identified yet. However, rapid development of the case reporting and surveillance system has lessened the impact of such bias in our study. Third, the results of our study are based upon the data which were collected from 2007 to 2011, which have merely reflected the association between HIV/AIDS cases among IDUs in China and macroscopic-factors during this period.

This work was supported by the National Scientific Research Mega-Project under the 12th Five-Year Plan of China (2012ZX10001001).

Correspondence should be addressed to: WANG Lu, Tel: 86-10-58900925, E-mail: wanglu64@163.com

Received: September 22, 2013;
Accepted: November 25, 2013

REFERENCES


