Introduction

Hepatitis C virus (HCV) is known as the major cause of chronic liver disease, human hepatic cirrhosis, and even hepatocellular carcinoma [1]. The fact that more and more HCV infections or patients who living with HCV has become a severe public health problem [2]. WHO has estimated that there are more than 150 million people with chronic liver disease which were caused by HCV infection in 2012, and over 350,000 people has been dead from HCV-related diseases worldwide each year [3]. Due to lack of vaccine and unavailability of effective therapy, the prevention of HCV infection has been a great challenge, especially in China, which is the largest developing country and owns one-fifths of the world’s population. The population of China had approached 1.35 billion at the end of 2011 [4]. Although the prevalence of HCV infection had been lower in China after mid-1990s and the overall prevalence of anti-HCV was less than 1% in recent years [5,6], the huge size of China population makes the absolute number of people infected with HCV enormous.

HCV infection is a public issue, which has an association with the interrelated social, geographic, historical and economic elements, just like other infectious diseases. Although various studies have already shown the prevalence of HCV infection varies with the different geographic positions in China [1,4,7], most of existing studies mainly focused on the spatial distribution of HCV genotypes and limited researches on the spatial distribution of HCV infection itself [8–10]. Shunquan Wu et al used spatial technique to work on the prevalence of HCV infection in Fujian province, and found that the HCV infection did have a relationship with spatial factors and the cluster of disease [11]. Apart from this study, the HCV infections in some provinces were
more serious than other surrounding provinces, which might attribute to poor economy, inadequate primary health infrastructure, or inappropriate treatments [1]. In addition, HCV infection shares the same transmission routes with HIV infection, which had been already observed cluster of disease in studies [12,13]. Especially for the intravenous drug users (IDUs), the prevalence of co-infection of HIV and HCV are high [14], and the HIV epidemics have association with geographic factors [15].

Since there is a geographic indicator related to the HIV infection, a macroscopic spatial analysis of HCV infection in this country is needed to probe whether the distribution of HCV infection was randomly distributed or not, which might facilitate prevention and control for HCV epidemic. In this study, we utilize the data of HCV reported cases in mainland of China from the year of 2005 to 2011 to explore the potential relationship between HCV infection and geographic distribution by using geographical information systems (GIS) and spatial statistics, which including general spatial autocorrelation (a tool was used to measure and analyze the degree of dependency among observations in a geographic space) and local spatial autocorrelation (a tool was used to detect cluster of high-value or low-value of observations).

Methods

Data collection and management

In 2004, after the Severe Acute Respiratory Syndrome (SARS) outbreak, China has established a world’s largest web-based disease reporting system, called China Information System for Disease Control and Prevention (CISDCP). Since then, all medical institutions should report diagnosed HCV infection cases to

![Figure 1.](image-url)
specific local Center for Disease Control and Prevention (CDC) through CISDCP. Demographic information (age, gender, address, registered residency places, time of onset of the disease etc.) were collected using standardized case report forms (CRFs) by conducting private interviews. After being checked, confirmed case reports were saved in the web-based system. The procedure of reporting HCV case through CISDCP is consistent around the whole country. This comprehensive and well-organized system, CISDCP, can provide both the number of HCV infection cases and the spatial distribution information of these cases across the country.

All HCV infection cases, identified during 2005 to 2011 in CISDCP, were included. In order to identify the residential locations of the reported HCV infection cases, the corresponding national standard geocodes at city level were included in the analysis. Any personal identifiers which might reveal the privacy of the participants were removed before data analysis in this study. Electronic maps were obtained from China CDC (CCDC). ArcGIS 10.1 software (ESRI Inc., Redlands, CA, USA) was used to create electronic maps and SPSS18.0 software (IBM Inc., Armonk, NY, USA) was used to process and analyze the data.

The data in this study was based on HCV regular monitoring system in China, and related ethics committee and Chinese government have approved for the system to collect patient data. Since this study focused on population-level analyses only and did not access any individually identifiable patient data, ethics committee approval was not particularly required in this study.

Anyone can apply for using the data in The data-center of China public health science (http://www.phsciencedata.cn/Share/en/data.jsp?id=9906073c-200a-4b44-8ffe-0867bfa42557) or email to data@chinacic.cn.

**Trend analysis**

We used Cochran-Armitage trend test to analyze the changing patterns of demographic and other characteristics, by ratios of gender (male/female), age group (≤30 years old/≤31 years old) and diagnosis type (clinical diagnosis/labatory diagnosis) of the identified HCV infection cases from 2005 to 2011. In the present study, α = 0.05 has been selected as the level of significance for the test.

**General spatial autocorrelation**

General spatial autocorrelation test statistic is a technique which is able to measure and analyze the spatial clusters in the data, and calculate the degree of dependency among observations in the whole geographic space [16]. In our study, general Moran’s Index was used to discover and measure the HCV infection clusters in

![Figure 2. Geographical distribution of the number of identified HCV infection cases reported at city level in the years 2005, 2007, 2009 and 2011 in China.](doi:10.1371/journal.pone.0110861.g002)
mainland China. The value of Moran’s Index was set between $[-1, 1]$. When the value of general Moran’s Index was $>0$ and Z-value $>1.96$, or the general Moran’s Index was $<0$ and Z-value $<-1.96$, it indicated that the distribution of identified HCV infection cases clustered in the whole area; otherwise, the distribution of the infection cases was random. The specific formula of general Moran’s Index was calculated as: [12]:

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

Where $n$ is the number of spatial units (cities), $x_i$ and $x_j$ were the observations from unit $i$ to unit $j$ about the phenomenon $X$; $w_{ij}$ represents the adjacent weight matrix. If the unit $i$ was adjacent to the unit $j$, then $w_{ij}$ would be 1; otherwise, it would be 0. In order to avoid the human influence of distance band of matrix, the threshold distance between cities was used in this study.

Local spatial autocorrelation

This method, also name local indicator of spatial association (LISA), was initially created to detect the clustering of cases of rare diseases [17]. The method focused on detecting specific local clusters of cases without any preconception about their locations. In other words, the aim of local spatial autocorrelation is to recognize clusters which may not be identifiable by general spatial autocorrelation. The Getis statistics was chosen as the parameter to identify the local clusters in the present study, and the formula of it was showed as below [12]:

$$G_i(d) = \frac{\sum_{i=1}^{n} \frac{n}{\sum_{j=1, j \neq i}^{n} w_{ij}} (x_i - \overline{x})x_j}{\sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} x_i}$$

The meaning of parameters are similar as general Moran’s Index formula, Z-test was conducted for the $G_i$ parameter. If Z-value was $>1.96$, the local clusters were identified as high-value correlations (statistically significant hotspot, meaning the city had a high number of cases and was surrounded by other cities with high number of cases as well); and if Z-value was $<-1.96$, the local clusters can be identified as low-value correlation (statistically significant coldspot, meaning the city had a low number of cases). We conducted local spatial autocorrelation at city level to detect the hotspots and coldspots of HCV infection in mainland China.

Results

Basic information

There were 774,787 identified HCV infection cases that had accurate geographic information during the study period (2005–2011). 7,568 cases (accounting for 0.98% of all identified) were excluded from the present study, due to lack of spatial information. Increased trend of number of identified HCV infection cases with years was observed (Figure 1A). During 2005–2011, the overall gender ratio (male/female) was 1.39; the mean age of all cases was 47.31 years (95%CI: 47.27 to 47.35) and, 16.51% of cases were ≤ 30 years. Almost 83.8% of HCV infection cases were identified by laboratory diagnosis and the rest were identified by clinical diagnosis.

Trend analysis

The ratio of gender (Male/Female) was generally decreasing, 1.53, 1.47, 1.43, 1.44, 1.39 and 1.35 from 2005 to 2011 respectively, Z-value = -18.53 ($P<0.001$). The ratio of age groups of cases ($\leq$30 years old/$\geq$31 years old) was 0.26 in 2005, while 0.16 in 2011. This trend also has statistical significance (Z-value = -51.03, $P<0.001$).

Clinically diagnosis case was defined as a patient was detected by clinically agencies when he/she seek for treatment, while laboratory diagnosis case was simply detected by laboratory. The ratio of diagnosis type (Clinical diagnosis/Laboratory diagnosis) was also presented decreasing trend during study period, from 0.47 in 2005 to 0.12 in 2011 (Z-value = -130.47, $P<0.001$). The trend analyses were described in Figure 1B.

Spatial analysis

The geographical distribution of identified HCV infection cases was found to be unbalanced in the mainland of China. As figure 2 presented, most cities have reported HCV infection cases during the study period, and the number of cities which have reported HCV infection cases were 339,341,341,342 in the year of 2005, 2007, 2009 and 2011, respectively. Henan province, Guangdong province, Guangxi Zhuang Autonomous Region, Xinjiang Uygur Autonomous Region and Jilin province have respectively reported more than 40,000 HCV infection cases during the period of 2005 to 2011. The total number of HCV identified cases in these five regions is 340,209, which accounted for 43.91% of all cases.

Firstly, general spatial autocorrelation was conducted for the cumulative number of HCV infection diagnosed between 2005

| Year  | Moran’s Index | Z-value | P-value |
|-------|---------------|---------|---------|
| 2005  | 0.123         | 9.497   | <0.01   |
| 2006  | 0.134         | 10.458  | <0.01   |
| 2007  | 0.155         | 11.983  | <0.01   |
| 2008  | 0.163         | 12.528  | <0.01   |
| 2009  | 0.180         | 13.771  | <0.01   |
| 2010  | 0.180         | 13.789  | <0.01   |
| 2011  | 0.203         | 15.460  | <0.01   |
| Total(2005–2011) | 0.178 | 13.655 | <0.01 |

doi:10.1371/journal.pone.0110861.t001

Spatial Analysis on HCV Infection in China
Figure 3. Hotspots and Coldspots of identified HCV infection cases at city level by year in China, 2005–2012.
doi:10.1371/journal.pone.0110861.g003
According to the previous study [18], the reported incidence of HCV infection in China was gradually increased during 1997 to 2011, especially since CISDCP was established in 2004. The upsurge was observed in both the prevalent number of HCV infection and its incidence indicating the emerging HCV infection epidemic in China.

Due to risky behaviors male conduct, for example, intravenous drug use and paid-blood/plasma donation, males are prone to get infected by HCV compared with female. However, the gender ratio (male/female) of HCV cases was decreasing in the present study, which was consistent with other studies [11,19]. The equality of women's status, for example, more chances to involve into social connections and continuous improvement of mobility might be the reason. The ratio of age groups (≤30 years old/≥31 years old) and diagnosis types (Clinical diagnosis/Laboratory diagnosis) both showed a decreasing trend.

In the present study, the proportion of people aged more than 30 years old was increasing by year, so older people became more likely to get infected by HCV. Persistent HCV infection risk factors caused by the cumulative effect (the number and degrees of risk factors growing with the HCV-infected patient getting older) and more opportunities to assess to the clinical invasive treatment might explain this phenomenon.

Much more infection cases were diagnosed by laboratory, indicating that large-scaled improvement of laboratory testing capabilities has promoted early detection to explore more cases than ever, which also contributed to the growing number of identified number of HCV cases. Instead, some hospitals might not have adequate capabilities to detect infection.

The number of cities which reported HCV infection cases remained stable by year, with the number of infection cases increasing in the same period. However, spatial analysis indicated that the hotspots and coldspots existed, which indicated the HCV infection and epidemic was not randomly distributed.

Henan province, Guangdong province, Guangxi Zhuang Autonomous Region, Xinjiang Uygur Autonomous Region and Jilin province accounted for more than 40 percent of identified HCV infection cases in China during the study period. The illegal blood/plasma donation, intravenous drug use and poorer economy status etc. might be attributed to HCV infection.
epidemic in these provinces. The northern and central areas reported more infection cases than the rest of regions in China during the study period. There are 11 provinces, municipalities or autonomous regions once had the hotspots located from 2005 to 2011, and the number of hotspots is decreasing by year. The central and border areas were the regions where the hotspots frequently located.

Owing to the poorer economy status, Henan province suffered from paid-blood/plasma donation since 1990s [20]. Although various interventions (voluntary donation, etc.) and blood donation law had been conducted for many years, the huge gap between demand and supplement motivates the blood mostly through employer-organized blood collection. However, the donors may not have been true volunteers, as they may be forced by the employer in some degrees [1]. In addition, the poor economy and low level of health status facilitated HCV infection in this area during the period (2005–2011).

Hotspots were observed in Guangdong province, Guangxi Zhuang Autonomous Region, and Xinjiang Uygur Region in most of 7 years. These provinces located in the heroin trafficking route, which begin from “Golden Triangle” and then go to the southwestern or west provinces in China [15]. The circulation of drug logically brings out some risky behaviors of HCV infection [21], for example, the needles or cottsins sharing, which were also the major reasons for HIV infection [22]. The northern hotspots frequently located in Changchun (Jilin province), Harbin (Heilongjiang province), both of two cities are the provincial capitals, which bring together many of medical resources and there were more likely to have clinical infection than other areas. In addition, Yanbian Korean Autonomous prefecture (Jilin province) was another location that a lot of case clustered, nationality and life style are the same, and the circulation of drug seems to be urgent. In this area during the period (2005–2011).

Spatial Analysis on HCV Infection in China

Acknowledgments

The authors thank all staff members responsible for HCV case reporting in medical institutes in mainland China, and all participants involving in this project for their efforts.

Author Contributions

Conceived and designed the experiments: Lu Wang JX. Performed the experiments: Lu Wang JX. Analyzed the data: Lu Wang JX. Contributed reagents/materials/analysis tools: Lu Wang JX RY LG QQ Liyan Wang ZD WG NW. Wrote the paper: Lu Wang JX FC.

References

1. Gao XF, Cui Q, Shi X, Su J, Peng ZH, et al. (2011) Prevalence and trend of hepatitis C virus infection among blood donors in Chinese mainland: a systematic review and meta-analysis. BMC Infect Dis 11: 88.
2. Shepard CW, Finelli L, Alter MJ (2005) Global epidemiology of hepatitis C virus infection. Lancet Infect Dis 5: 558–567.
3. World Health Organization (July 2013) Hepatitis C fact sheet. Available: http://www.who.int/mediacentre/factsheets/fs164/en. Accessed 26 March 2014.
4. National Bureau of Statistics of China (2011). China statistical yearbook 2011. Beijing Statistics Press.
5. Cui Y, Jia JD (2013) Update on epidemiology of hepatitis B and C in China. J Gastroenterol Hepatoll 28 Suppl 1: 7–10.
6. Kuang YQ, Yan J, Li Y, Huang X, Wang Y, et al. (2013) Molecular epidemiologic characterization of a clustering HCV infection caused by inappropriate medical care in Heyuan city of Guangdong, China. PLoS One 8: e52304.
7. Li D, Long Y, Wang T, Xiao D, Zhang J, et al. (2013) Epidemiology of hepatitis C virus infection in highly endemic HBV areas in China. PLoS One 8: e54153.
8. Zeng G, Wang Z, Wen S, Jiang J, Wang L, et al. (2005) Geographic distribution, virologic and clinical characteristics of hepatitis B virus genotypes in China. J Viral Hepat 12: 609–617.
9. Li L, Nakata T, He Y, Fu Y, Hagedorn CH, et al. (2005) Hepatitis C virus genotype distribution in China: predominance of closely related subtype 1b isolates and existence of new genotype 6 variants. J Med Virol 75: 539–549.
10. Xia XS, Lu L, Tee KK, Zhao WH, Wu JG, et al. (2008) The unique HCV genotype distribution and the discovery of a novel subtype 6u among IDUs co-infected with HIV-1 in Yunnan, China. J Med Virol 80: 1142–1152.
11. Wu SQ, Wu FQ, Hong RT, He J (2012) Incidence analyses and space-time cluster detection of hepatitis C in Fujian province of China from 2006 to 2010. PLoS One 7: e46872.
12. Peng ZH, Cheng YJ, Reilly KH, Wang L, Qin QQ, et al. (2011) Spatial distribution of HIV/AIDS in Yunnan province, People's Republic of China. Geospat Health 5: 177–182.
13. Xia XS, Lu L, Chen SY, Li SY, Zhang MX, et al. (2011) Tracking the evolution of HIV/AIDS in China from 1989-2009 to inform future prevention and control efforts. PLoS One 6: e25671.
14. Chen X, He JM, Ding LS, Zhang QQ, Zou XB, et al. (2013) Prevalence of hepatitis B virus and hepatitis C virus in patients with human immunodeficiency virus infection in Central China. Arch Virol 158: 1889–1894.
15. Beyer C, Razak MH, Lisam K, Chen J, Liu W, et al. (2000) Overland heroin trafficking routes and HIV-1 spread in south and south-east Asia. AIDS 14: 75–83.
16. Tobler WR (1970) A computer movie simulating urban growth in the Detroit area. Economic geography 46: 234–249.
17. Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geographical analysis 27: 290–300.
18. Hajinazadeh B, Greybel J, Dore GJ (2013) Epidemiology and natural history of HCV infection. Nat Rev Gastroenterol Hepatol 10: 553–562.
19. Qin QQ, Guo W, Wang LY, Yan RX, Ge L, et al. (2013) Epidemiological characteristics of hepatitis C in China, 1997–2011. Chinese J Epidemiol 34: 549–551. [In Chinese].
20. Tao D, Li L, Liu Y, Li H, Xu X, et al. (2012) Different HCV genotype distributions of HIV-infected individuals in Henan and Guangxi, China. PLoS One 7: e50343.
21. Williams GT, Liu W, Levy JA (2011) Crossing Over: Drug Network Characteristics and injection risk along the China-Myanmar border. AIDS Behav 15: 1011–1016.
22. Choi SY, Cheung YW, Jiang ZQ (2007) Ethnicity and risk factors in needle sharing among intravenous drug users in Sichuan province, China. AIDS Care 19: 1–8.

23. Huang B, Deng LQ, Lu XR (2013) Analysis on influencing factors of HCV infection among the population in Yanbian prefecture. Practical Preventive Medicine 20: 556–557. [In Chinese].