Geographical Clustering Analysis of Birth Defects in Guangxi

Zhenren Peng

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Jie Wei

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Jie Qin

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Biyan Chen

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Xiuning Huang

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Pengshu Song

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Lifang Liang

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Xiaoxia Qiu

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Junrong Liang

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Hongwei Wei

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Sheng He (✉ heshengbiol@163.com)

The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region

Research Article

Keywords: Birth defects (BD), Guangxi Zhuang, spatial clustering patterns, geographical clustering of BD

Posted Date: July 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-737750/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Geographical clustering analysis of birth defects in Guangxi

Zhenren Peng, Jie Wei, Jie Qin, Biyan Chen, Xiuning Huang, Pengshu Song, Lifang Liang, Xiaoxia Qiu, Junrong Liang, Hongwei Wei*, and Sheng He*

Key Laboratory of Basic Research for Guangxi Birth Defects Control and Prevention, Department of Research Center for Guangxi Birth Defects Control and Prevention, The Maternal & Child Health Hospital of Guangxi Zhuang Autonomous Region, Nanning, 530000, Guangxi Zhuang Autonomous Region, China.

*corresponding author email: weihongwei1965@163.com; heshengbiol@163.com

Abstract

Birth defects (BD) is a big public health issue in Guangxi Zhuang Autonomous Region of China. The overall prevalence of BD in Guangxi is about 1% and higher than most other provinces of China. However, the geographical clustering variations in BD of Guangxi has not been described. Therefore, the aim of this study was to explore and detect the spatial clustering patterns of BD prevalence across a well-defined geographic space. The data were obtained from Guangxi birth defects monitoring network (GXBDMN) from 2016 to 2020, which collected socio-demographic and clinical information from perinatal infants between 28 weeks of gestation and 7 days postnatal. The spatial autocorrelation analysis and hot spot analysis will be used to explore the geographical clustering of BD prevalence in 70 counties and 41 districts of Guangxi in this study. A total of 44,418 perinatal infants were born with BD from 2016 to 2020. The overall prevalence of BD was 122.47/10,000 [95% confidence interval (CI): 121.34-123.60/10,000]. The local indicators of spatial association (LISA) statistic and G_i^* statistic showed that the spatial clustering patterns of BD prevalence changed over time, and the largest High-High clustering area and hot spot area were both identified in the city of Nanning. Therefore, the spatial clustering patterns of BD prevalence in Guangxi is very significant. Spatial cluster analysis can provide reliable and accurate spatial distribution patterns in BD control and prevention.

Introduction

Birth defects (BD), also known as congenital anomalies, refer to any structural or functional abnormalities that occur during intrauterine, including metabolic disorders¹⁻³. BD is one of the significant causes of spontaneous abortions, stillbirths, and also death and disability among infants and children under 5 years old³. It is estimated that about 3-6% (about 7.9 million) of infants suffer from serious BD and more than 3.3 million infants and children die of BD every year in the world, and also BD is one of the main reasons for the loss of disability adjusted life years (DALYs) of infants aged 0-1⁴⁻⁷. The prevalence of BD in China is about 5.6%, which means around 0.9 million infants were born with BD every year, which has a serious impact on the survival and quality of life of infants and brings great financial burden to their families^{8,9}.

Spatial epidemiology is an emerging (or re-emerging) interdisciplinary which is based on geographic information system (GIS) spatial analysis technology, can be used to describe and analyze the risk factors of demographics, environment, behavior, socio-economics, genetics and infectious disease^{10,11}. In recent years, spatial epidemiology has been widely used in the study of the relationship between environment and health, and plays a very significant role in the field of public health^{12,13}. It has shown that the prevalence of BD is closely related to geographical location^{14,15}. The prevalence of BD in Guangxi is about 1%, higher than most other provinces of China, and varies greatly in different regions^{16,17}. In this study, we applied the spatial autocorrelation analysis and hot spot analysis in

the geographical clustering of BD prevalence in 70 counties and 41 districts of Guangxi. It will provide scientific basis for the future development of BD prevention and control strategies upon this result.

Results

BD prevalence mapping

The prevalence of BD in the 70 counties and 41 districts of Guangxi was mapping and shown in Figure 1, in which the highest values of BD prevalence in the year of 2016 to 2020 were 294.79/10,000 (Chengzhong district of Liuzhou city), 266.28/10,000 (Jinchengjiang district of Hechi city), 408.60/10,000 (Xingning district of Nanning city), 522.88/10,000 (Xingning district of Nanning city), and 636.58/10,000 (Xingning district of Nanning city), respectively.

Spatial autocorrelation analysis of BD prevalence

The Z test of global Moran's I statistic showed that spatial autocorrelation was significant at the 95% confidence interval (*CI*) level, and the global Moran's I index in the year of 2016 to 2020 were 0.11, 0.19, 0.28, 0.33, 0.36, respectively. These results suggest positive spatial autocorrelation of BD prevalence in the entire study area.

The local indicators of spatial association (LISA) statistic showed that High-High cluster, High-Low cluster, and Low-High cluster were the significant local spatial clustering patterns of BD prevalence in the study area from 2016 to 2020, and the spatial clustering patterns was illustrated in Figure 2. The above three significant local spatial clustering patterns were all shown in the year 2017, and High-High cluster was shown in the year of 2016 to 2020 but the year of 2019 and 2020 were only shown High-High cluster, and High-Low cluster was only shown in the year of 2016 and 2017, and Low-High cluster was only shown in the year of 2017 and 2018. The largest High-High clustering area located in the city of Nanning of the year 2020, which were Xingning, Qingxiu, Jiangnan, Xixiangtang, Liangqing and Wuming districts, and Binyang county. The second largest High-High clustering area located in the city of Nanning of the year 2017 to 2019, which were Xingning, Qingxiu, Jiangnan, Xixiangtang and Wuming districts, and Binyang county. And the third largest High-High clustering area located in the city of Nanning of the year 2016, which were Xingning, Xixiangtang and Wuming districts.

Hot spots analysis of BD prevalence

G_i^* statistic showed that spatial clustering was significant at the 90% *CI* level of BD prevalence in this study area. Hot spot area and cold spot area were both shown in the study area of the year 2016 to 2018, but only hot spot area was shown in the year 2019 and 2020. The largest hot spot area located in the cities of Nanning and Chongzuo of the year 2018 to 2020, which were Xingning, Qingxiu, Jiangnan, Xixiangtang, Liangqing, Yongning and Wuming districts, and long'an, Shanglin, Binyang and Fusui counties. The second largest hot spot area located in Nanning city of the year 2017, which were Xingning, Qingxiu, Jiangnan, Xixiangtang, Liangqing, Yongning and Wuming districts, and long'an, Shanglin and Binyang counties, but another large hot spot area located in Liuzhou city which were Chengzhong, Yufeng, Liunan, Liubei and Liujiang districts, and Liucheng and Luzhai counties. The largest hot spot area located in Liuzhou city of the year 2016, which were Chengzhong, Yufeng, Liunan, Liubei and Liujiang districts, and Liucheng and Luzhai counties, but another large hot spot area located in Nanning city which were Xingning, Qingxiu, Xixiangtang and Wuming districts. In addition, the largest cold spot area located in the city of Beihai and Yulin of the year 2016, which were Haicheng, Yinhai and Tieshangang districts, and Hepu and Bobai counties. And the second largest cold spot area located in the city of Baise and Hechi of the year 2017, which were Leye, Tian'e and Fengshan counties. The results of G_i^* statistic were all shown in Figure 3.

Discussion

Spatial epidemiology based on GIS spatial analysis technology plays an important role in control and prevention of diseases¹³, which usually be applied in disease surveillance and health risk analysis in public health^{18,19}. Although spatial epidemiology has been applied to explore BD and maternal health problems in recent years²⁰⁻²⁴, but most studies were only applied simple descriptive mapping of spatial distribution patterns²⁵, especially in Guangxi, the spatial epidemiology of BD is not well understood. This study is the first spatial epidemiology report on BD prevalence of Guangxi by using GIS technology.

Spatial cluster analysis is a uniquely interdisciplinary research, and so it is very important to exchange ideas among applied epidemiology researchers and spatial statisticians²⁶. Spatial cluster analysis is usually used to analyze the spatial distribution patterns for BD prevalence²⁷⁻²⁹. In our study, global Moran's I statistic, LISA and G_i^* statistic were used to explore and detect the spatial clustering pattern of BD prevalence in Guangxi from 2016 to 2020.

The LISA statistic results showed that the spatial clustering pattern of BD prevalence changed over time. The number of counties and districts contained in the largest High-High clustering area increased from 3 in 2016 to 6 after 2016 and then increased to 7 in 2020, and the High-Low and Low-High clustering patterns disappeared after 2018, but the largest High-High clustering area were all located in the city of Nanning from 2016 to 2020.

The G_i^* statistic results showed that the spatial clustering pattern of BD prevalence also changed over time. The number of counties and districts contained in the largest hot spot area increased from 7 in 2016 to 11 after 2017. The cold spot area disappeared after 2018, but the location of the largest hot spot area changed from Liuzhou city in 2016 to Nanning city in 2017 and then changed to the cities of Nanning and Chongzuo after 2017. Nevertheless, the hot spot area located in Chongzuo city was only the Fusui county, which it is adjacent to Nanning city. And the city of Nanning was also a large hot spot area in 2016.

So, these spatial cluster analysis results indicated that a significant spatial clustering pattern in Guangxi from 2016 to 2020, it means that there was a clustering of high values of BD prevalence in the city of Nanning. And also, these spatial cluster analysis methods explored spatial clustering of BD prevalence provided strong evidence for researchers. Spatial clustering of BD prevalence was centered in the city of Nanning, which is known for the provincial capital city of Guangxi in China, and its gross domestic product (GDP) is the highest in Guangxi. Moreover, Nanning provides professional pre/postnatal health care service. For these reasons, more pregnant women chose to be hospitalized in the hospital which has high professional maternal and child health care level, especially those who have severe problems found in other hospitals. As a result of this, the defects prevalence of perinatal infants may be higher in the city of Nanning. This suggests that it may be an important problem and need to be considered in the future when health policy formulated by our health administrative departments.

However, our study has two major limitations. First, the calculation of BD prevalence was a critical process in this study. It is widely known that the importance of population-based cohort study in precision medicine and translational medicine. Unlike a population-based cohort study, the data of our study were collected by Guangxi birth defects monitoring network (GXBDMN), which is a hospital-based, passive surveillance system. So, it may not fully reflect the actual BD prevalence of each county and district in Guangxi.

The second limitation was that the related risk factors of BD prevalence were not included to analyze in this study. The geographical variations of BD prevalence may be explained by various factors, such as socio-economic development level, environmental factors such air and water pollutions, the medical level of maternal and child health care institutions, and lifestyle and behavior of pregnant women. The main aim of this study in the paper was to detect spatial clustering patterns and identify the location of spatial cluster areas. So, the related risk factors of BD prevalence were not included to analyze in this study. We hope that more insight into the epidemiology of BD by using other spatial analysis methods such spatial regression analysis may be considered in our future study based on this research.

In conclusion, the spatial clustering patterns of BD prevalence in Guangxi is very significant. Spatial cluster

analysis can provide reliable and accurate spatial distribution patterns in BD control and prevention. Especially LISA statistic and G_i^* statistic enable us to explore and detect more reliably location of spatial clustering area when researchers have some prior knowledge of spatial epidemiology and GIS spatial analysis technology.

Methods

Study design

In this study, the prevalence of BD of perinatal infants between 28 weeks of gestation and 7 days postnatal were mapped at the county level. Then spatial autocorrelation analysis and hot spot analysis were used to detect the spatial cluster type of the prevalence of BD by using GIS spatial analysis technology. Finally, all data analysis, including spatial analysis, mapping and data preprocessing were conducted using ArcGIS 10.2, which is a GIS software and developed by Esri company. And all the spatial analysis results were mapped by using ArcGIS 10.2.

Study area

The study area Guangxi Zhuang Autonomous Region is located in southwestern China (see Figure 4), which includes 70 counties and 41 districts. With a population of approximately 50 million, and the population density is about 210 persons/km². The basic geographical data is the county and district boundaries of Guangxi were provided in the form of shapefile and downloaded from the National Platform for Common Geospatial Information Services.

Data source

Data were collected by GXBDMN from 2016 to 2020. The perinatal infants between 28 weeks of gestation and 7 days postnatal were monitored by GXBDMN. The diagnosis of BD for perinatal infants were according to the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10), and there were more than 40 types of BD in our study (Table 1). In this study, there were 3,626,871 cases were monitored by GXBDMN, among which 44,418 BD cases were reported. The overall prevalence of BD was 122.47/10,000 [95% *CI*: 121.34-123.60/10,000].

Data preprocessing

The prevalence of BD of 70 counties and 41 districts were geocoded by regionalism code in the shapefile of Guangxi which comprising of county and district boundaries.

Spatial cluster analysis

Spatial autocorrelation analysis and hot spot analysis are two important analytical processes for geographical clustering analyses. Generally, spatial autocorrelation has two analytical methods: global spatial autocorrelation³⁰ and local spatial autocorrelation³¹. In our study, BD prevalence of Guangxi in the year of 2016 to 2020 were analyzed by using spatial autocorrelation analysis and hot spot analysis.

Global spatial autocorrelation is a correlation analysis of the entire study area which assumes that all the spatial elements (stochastic variables) are on a plane³⁰. The function of global spatial autocorrelation is used to describe the overall spatial distribution of a phenomenon and judge whether the phenomenon has spatial cluster area in the entire study area. Global Moran's I statistic³⁰ is usually be used to test for spatial cluster in global spatial autocorrelation analysis.

The global Moran's I statistic is the stochastic processes of identifying stochastic phenomena which are distributed in space of two dimensions and can evaluate the pattern of spatial cluster (spatial clustering pattern, discrete pattern or random pattern)³⁰. In our study, global Moran's I statistic was firstly applied to explore the spatial clustering pattern for the prevalence of BD in Guangxi. The formula of the global Moran's I statistic is:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{\sum_{i=1}^n z_i^2}$$

$$= \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{i,j}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} z_i z_j}{\sum_{i=1}^n z_i^2} \quad \text{equation (1)}$$

In equation (1), where z_i is the deviation between the attribute of unit i and its average value ($x_i - \bar{X}$), $w_{i,j}$ is the spatial weight between units i and j , n is the number of observations (units), and S_0 is the aggregation of all spatial weights.

The Z test usually using to test global Moran's I index. The values of Global Moran's I index are ranging from -1 to +1. When Z scores in which values > 1.96 or < -1.96 show that spatial autocorrelation is significant at the 95% CI level. So, when P value < 0.05 and global Moran's I index > 0 indicates positive spatial autocorrelation, it means high value is adjacent to high value or low value is adjacent to low value; when P value < 0.05 and global Moran's I index < 0 indicates negative spatial autocorrelation, it means low value is adjacent to high value; and when P value > 0.05 and global Moran's I index is 0 indicates a random spatial pattern.

Local spatial autocorrelation analysis is advanced by Luc Amelin³¹, in his paper, a new general class of LISA statistic is used to decompose of global Moran's I statistic and into the contribution of each observation. LISA statistic is generally applied to study local spatial clustering when global Moran's I statistic is significant, and it not only may be interpreted as indicators of local pockets of nonstationarity or hot spots, but also may be used to evaluate the influence of individual locations on the magnitude of the global Moran's I statistic and to identify outliers³¹. The formula of the LISA statistic is:

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^n w_{i,j} (x_i - \bar{X})$$

$$= \frac{x_i - \bar{X}}{\left(\frac{\sum_{j=1, j \neq i}^n (x_i - \bar{X})^2}{n-1} - \bar{X}^2 \right)} \sum_{j=1, j \neq i}^n w_{i,j} (x_i - \bar{X})$$

equation (2)

In equation (2), where x_i is the spatial attribute value of unit i , $w_{i,j}$ is the spatial weight between units i and j , n is the number of observations (units), and S_i^2 is the aggregation of all spatial weights.

The Z test is also used to test LISA statistic. When Z scores in which values > 1.96 or < -1.96 show that spatial autocorrelation is significant at the 95% CI level. So, the local spatial clustering patterns can be classified 4 types when P value < 0.05 : High-High cluster (clustering of high values), Low-Low cluster (clustering of low values), High-Low cluster (outliers in which a high value is surrounded primarily by low values), and Low-High cluster (outliers in which a low value is surrounded primarily by high values)³¹.

Hot spot analysis, also known as G_i^* statistic, is introduced and developed by Getis and Ord³², which can be used as a method of measuring of spatial association in a number of circumstances and can enable researchers to explore local "pockets" (hot or cold spots) of dependence that may not show up when using global Moran's I statistic³². The formula of the G_i^* statistic is:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}}$$

$$= \frac{\sum_{j=1}^n x_j}{n}$$

S

$$= \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2}$$

In equation (3), (4) and (5), where x_j is the spatial attribute value of unit j , w_{ij} is the spatial weight between units i and j , n is the number of observations (units), S_i^2 is the aggregation of all spatial weights, $\bar{X}$ is the mean value of all attribute values in the study area, and S is the standard deviation (SD) of all attribute values in the study area.

The Z test is also used to test G_i^* statistic. When Z scores in which values > 1.65 or < -1.65 show that spatial clustering is significant at the 90% CI level. So, when P value < 0.10 and Z score > 0 indicates the study area has hot spots; when P value < 0.10 and Z score < 0 indicates the study area has cold spots; and when P value > 0.10 indicates the study area has not hot and cold spots.

References

1. Groisman, B. *et al.* Join World Birth Defects Day. *Pediatr. Res.* **86**, 3-4. <https://doi.org/10.1038/s41390-019-0392-x> (2019).
2. Bale, J. R., Stoll, B. J. & Lucas, A. O. Reducing Birth Defects: Meeting the Challenge in the Developing World. (ed. Committee on Improving Birth Outcomes) 1-253 (National Academy of Sciences, 2003).
3. World Health Organization. in Sixty-third World Health Assembly. 1-7 (WHO, 2010).
4. Centers for Disease Control and Prevention. World Birth Defects Day. <https://www.cdc.gov/ncbddd/birthdefects/features/birth-defects-day.html> (2020).
5. Howse, J. L., Howson, C. P. & Katz, M. Reducing the global toll of birth defects. *Lancet* **365**, 1846-1847. [https://doi.org/10.1016/s0140-6736\(05\)66611-1](https://doi.org/10.1016/s0140-6736(05)66611-1) (2005).
6. Christianson, A., Howson, C. P. & Modell, B. March of Dimes Global Report on Birth Defects: The Hidden Toll of Dying and Disabled Children. <https://www.marchofdimes.org/materials/global-report-on-birth-defects-the-hidden-toll-of-dying-and-disabled-children-executive-summary.pdf> (2006).
7. Collaborators, G. D. a. H. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1603-1658. [https://doi.org/10.1016/s0140-6736\(16\)31460-x](https://doi.org/10.1016/s0140-6736(16)31460-x) (2016).
8. National Health Commission of the people's Republic of China. Notice on printing and distributing the national program for comprehensive prevention and control of birth defects (GWBFYF [2018] No.19). <http://www.nhc.gov.cn/fys/s3589/201812/9644ce7d265342779099d54b6962a4e0.shtml> (2018).
9. Ministry of Health of the People's Republic of China. Report on prevention and treatment of birth defects in China (2012). <http://www.gov.cn/gzdt/att/att/site1/20120912/1c6f6506c7f811bacf9301.pdf> (2012).
10. Elliott, P. & Wartenberg, D. Spatial epidemiology: current approaches and future challenges. *Environ Health Perspect* **112**, 998-1006. <https://doi.org/10.1289/ehp.6735> (2004).
11. Ostfeld, R. S., Glass, G. E. & Keesing, F. Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends Ecol Evol* **20**, 328-336. <https://doi.org/10.1016/j.tree.2005.03.009> (2005).
12. Lee, J., Kamenetsky, M. E., Gangnon, R. E. & Zhu, J. Clustered spatio-temporal varying coefficient regression model. *Stat. Med.* **40**, 465-480. <https://doi.org/10.1002/sim.8785> (2021).
13. Liu, K., Li, X. L. & Shao, Z. J. Spatial epidemiology plays an important role in control and prevention of diseases. *Chin. J. Epidemiol.* **39**, 1143-1145. <https://doi.org/10.3760/cma.j.issn.0254-6450.2018.09.001> (2018).
14. Liao, Y. *et al.* Temporal and Spatial Analysis of Neural Tube Defects and Detection of Geographical Factors in Shanxi Province, China. *PLoS One* **11**, e0150332. <https://doi.org/10.1371/journal.pone.0150332> (2016).
15. Liu, Y. *et al.* Geographical variations in maternal lifestyles during pregnancy associated with congenital heart defects among live births in Shaanxi province, Northwestern China. *Sci. Rep.* **10**, 12958. <https://doi.org/10.1038/s41598-020-69788-0> (2020).

16. Health Commission of Guangxi Zhuang Autonomous Region. Notice of the health commission of Guangxi Autonomous Region on printing and distributing the comprehensive prevention and control program for birth defects in Guangxi (GWFYF [2019] No.7). http://wsjkw.gxzf.gov.cn/xxgk_49493/fdzdgk/wsjszh/fybj/t5236305.shtml (2018).
17. Dong, B. Q. *et al.* Study on the characteristics of major birth defects in 1.69 million cases of fetus in Guangxi Zhuang Autonomous Region. *Chin. J. Epidemiol.* **40**, 1554-1559. <https://doi.org/10.3760/cma.j.issn.0254-6450.2019.12.009> (2019).
18. Shaw, N. & McGuire, S. Understanding the use of geographical information systems (GIS) in health informatics research: A review. *Journal of innovation in health informatics* **24**, 228-233. <https://doi.org/10.14236/jhi.v24i2.940> (2017).
19. Kirby, R. S., Delmelle, E. & Eberth, J. M. Advances in spatial epidemiology and geographic information systems. *Ann. Epidemiol.* **27**, 1-9. <https://doi.org/10.1016/j.annepidem.2016.12.001> (2017).
20. Zhang, N., Liao, Y. & Ren, Z. Evaluating an intervention for neural tube defects in coal mining cites in China: a temporal and spatial analysis. *International health* **13**, 161-169. <https://doi.org/10.1093/inthealth/ihaa035> (2021).
21. Zhao, L. *et al.* Birth prevalence of congenital heart disease in China, 1980-2019: a systematic review and meta-analysis of 617 studies. *Eur. J. Epidemiol.* **35**, 631-642. <https://doi.org/10.1007/s10654-020-00653-0> (2020).
22. Meng, J., Zhang, Y. & Zheng, Q. Spatial epidemiology in perinatal medicine. *Chin J Perinat Med* **23**, 272-275. <https://doi.org/10.3760/cma.j.cn113903-20190918-00571> (2020).
23. Hansen, S. P., Messer, T. L. & Mittelstet, A. R. Mitigating the risk of atrazine exposure: Identifying hot spots and hot times in surface waters across Nebraska, USA. *J Environ Manage* **250**, 109424. <https://doi.org/10.1016/j.jenvman.2019.109424> (2019).
24. Lantos, P. M., Permar, S. R., Hoffman, K. & Swamy, G. K. The Excess Burden of Cytomegalovirus in African American Communities: A Geospatial Analysis. *Open forum infectious diseases* **2**, 1-7. <https://doi.org/10.1093/ofid/ofv180> (2015).
25. Ahmadian, L., Salehi, F. & Bahaadinbeigy, K. Application of geographic information systems in maternal health: a scoping review. *East. Mediterr. Health Journal* **26**, 1403-1414. <https://doi.org/10.26719/emhj.20.095> (2020).
26. Fritz, C. E., Schuurman, N., Robertson, C. & Lear, S. A scoping review of spatial cluster analysis techniques for point-event data. *Geospatial health* **7**, 183-198. <https://doi.org/10.4081/gh.2013.79> (2013).
27. Ma, L. G. *et al.* Spatial patterns of the congenital heart disease prevalence among 0- to 14-year-old children in Sichuan Basin, P. R China, from 2004 to 2009. *BMC Public Health* **14**, 595. <https://doi.org/10.1186/1471-2458-14-595> (2014).
28. Avilés, L. A., Alvelo-Maldonado, L., Padró-Mojica, I., Seguinot, J. & Jorge, J. C. Risk factors, prevalence trend, and clustering of hypospadias cases in Puerto Rico. *J. Pediatr. Urol.* **10**, 1076-1082. <https://doi.org/10.1016/j.jpuro.2014.03.014> (2014).
29. Winston, J. J., Meyer, R. E. & Emch, M. E. Geographic analysis of individual and environmental risk factors for hypospadias births. *Birth Defects Res. A Clin. Mol. Teratol.* **100**, 887-894. <https://doi.org/10.1002/bdra.23306> (2014).
30. Moran, P. A. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17-23. <https://doi.org/10.1093/biomet/37.1-2.17> (1950).
31. Amelin, L. Local Indicators of Spatial Association-LISA. *Geographical Analysis* **27**, 93-115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x> (1995).
32. Getis, A. & Ord, J. K. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis* **24**, 189-206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x> (1992).

Acknowledgements

Not applicable.

Author contributions

Z.P. designed the study, analyzed the data and wrote the manuscript text. J.W., X.H., P.S. and L.L. assisted to collect the data. J.Q., B.C., X.Q., J.L., H.W. and S.H. provided critical comments in reviewing and revising the manuscript. All authors have read and approved the manuscript.

Funding

This study was supported by Self Funded Program of Guangxi Health Commission (grant number: Z20210056),
Appropriate Technology Development and Application Program of Guangxi Medicine and Health (grant numbers:
S201613 and S201614), and The National Key Research and Development Program of China (grant numbers:
2018YFC1002501 and 2019YFC012190)

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.W. and S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and
institutional affiliations.

Figures

Figure 1. BD prevalence map of the year 2016 to 2020

Figure 2. LISA map of BD prevalence in the year 2016 to 2020

Figure 3. Hot spots analysis map of BD prevalence in the year 2016 to 2020

Figure 4. Study area's geographical location map

Tables

Birth defects and ICD-10 code

(1) Nervous system

Anencephaly (Q00); Encephalocele (Q01); Microcephaly (Q02); Congenital hydrocephalus (Q03); Spina bifida (Q05)

(2) Eyes, ears, faces and neck

Congenital cataracts (Q12.0); Anophthalmia, Microphthalmia, Megalophthalmos (Q11); Congenital anotia (Q16.0); Microtia (Q17.2); Accessory auricle (Q17.0)

(3) Circulatory system

Ventricular septal defects (Q21.0); Atrial septal defects (Q21.1); Tetralogy of Fallot (Q21.3); Hypoplastic left heart syndrome (Q23.4); Patent ductus arteriosus (Q25.0); Transposition of the great vessels (Q20.3); Pulmonary artery stenosis (Q25.6)

(4) Cleft lip and palate

Cleft palate (Q35); Cleft lip (Q36); Cleft palate with cleft lip (Q37)

(5) Digestive system

Esophageal stenosis and atresia (Q39.0, Q39.1, Q39.2, Q39.3); Anorectal stenosis and atresia (Q42.0, Q42.1, Q42.2, Q42.3)

(6) Reproductive system

Undescended testis (Q53); Hypospadias (Q54); Ambiguous genitalia and pseudohermaphroditism (Q56)

(7) Urinary system

Bladder exstrophy (Q64.1)

(8) Locomotor system

Talipes equinovarus (Q66.0); Polydactyly (Q69); Syndactyly (Q70); (Q71); (Q72); Congenital diaphragmatic hernia (Q79.0); Congenital omphalocele (Q79.2); Gastroschisis (Q79.3)

(9) Other congenital anomalies

Conjoined twins (Q89.4)

(10) Chromosome abnormalities

Down syndrome (Q90); 18-Trisomy syndrome (Q91); Hemangioma and lymphangioma (D18);

Thalassemia (D56)

(11) Other congenital abnormalities

Inguinal hernia (K40); Umbilical hernia (K42); Teratoma (D48.9); Congenital phenylketonuria (E70.0);

Congenital hypothyroidism (E03)

325

Table 1. Birth defects eligible for inclusion in the GXBDMN