

中国湖区五省血吸虫病聚集区域地理信息系统综合探测分析

赵飞 朱蓉 张丽娟 张志杰 李源培 何明祯 周艺彪 郭家钢 赵根明 姜庆五

【摘要】 目的 综合探测2008年中国湖区5省血吸虫病传播的聚集区域,为血吸虫病防治提供依据,也为相关疾病聚集区域探测方法的建立提供参考。**方法** 基于县级地理信息系统(GIS)基础上构建空间数据库,运用全局空间自相关Moran's I、局部空间自相关Getis-Ord Gi与SaTScan软件综合探测血吸虫病聚集区域,并对结果进行比较。**结果** 全局空间自相关结果提示从整体研究区域来看,血吸虫病感染率分布呈现空间聚集性($P<0.05$);局部空间自相关结果提示有50个县(区)的局部空间自相关结果有统计学意义($P<0.05$),且 z 值均 >0 ,提示存在高值聚集;SaTScan统计结果探测出5个聚集区域,与局部空间自相关的结果较为吻合。**结论** 湖区5省沿江地区仍然是中国血吸虫病防治的重点,尤以湖北、湖南两省交界地带空间聚集性最高、范围最大。

【关键词】 血吸虫病;地理信息系统;聚集区域

Integrated detection and analysis on the clusters of schistosomiasis based on geographic information system ZHAO Fei¹, ZHU Rong², ZHANG Li-juan², ZHANG Zhi-jie¹, LI Yuan-pei¹, HE Ming-zhen¹, ZHOU Yi-biao¹, GUO Jia-gang², ZHAO Gen-ming¹, JIANG Qing-wu¹. 1 Department of Epidemiology, School of Public Health, Fudan University, Key Laboratory on Public Health Safety, Ministry of Education, Shanghai 200032, China; 2 National Institute of Parasitic Disease, Chinese Center for Disease Control and Prevention

Corresponding author: JIANG Qing-wu, Email: jiangqw@fudan.edu.cn

This work was supported by grants from the National Natural Science Foundation of China (No. 30590374), the National Important Technologies Project (No. 2008ZX10004-011), Construction of Key Discipline of Shanghai Municipality (No. B118) and the Foundation of Science Innovation of Key Discipline, Fudan University.

【Abstract】 Objective To detect and analyze the clusters of schistosomiasis on marshland and lake areas based on geographic information system (GIS) in 2008 and to provide suggestions for the development of integrated methodology on the detection of clusters on related diseases. **Methods** Moran's I of global spatial autocorrelation, Getis-Ord Gi of local spatial autocorrelation and SaTScan were used to detect the schistosomiasis clusters based on GIS and comparison of the results for different methods were performed. **Results** Results from the global Moran's I tests for all the marshland and lake areas related to the schistosomiasis were statistically significant ($P<0.05$) and indicated spatial heterogeneity; the z -score outcomes as calculated by Getis-Ord Gi indicated high cluster that 50 clusters were categorized at the 0.05 significance level and the z -score of these 45 clusters were more than 0. Results of SaTScan statistics appeared the same as local spatial autocorrelation and almost showing the existence of 5 cluster areas. **Conclusion** The geographical distribution of clusters where schistosomiasis was prevalent showed that it was closely corresponded to the middle and lower Yangtse river and in particular, many clusters were located near the boundary of Hubei and Hunan province.

【Key words】 Schistosomiasis; Geographic information system; Cluster areas

DOI:10.3760/cma.j.issn.0254-6450.2010.11.016

基金项目:国家自然科学基金(30590374);国家科技重大专项(2008ZX10004-011);上海市重点学科建设项目(B118);复旦大学重点学科创新人才培养计划基金

作者单位:200032 上海,复旦大学公共卫生学院流行病学教研室 公共卫生安全教育部重点实验室(赵飞、张志杰、李源培、何明祯、周艺彪、赵根明、姜庆五);中国疾病预防控制中心寄生虫病预防控制所(朱蓉、张丽娟、郭家钢)

通信作者:姜庆五, Email:jiangqw@fudan.edu.cn

目前我国血吸虫病疫区主要集中在江苏、安徽、江西、湖北、湖南 5 个湖区省和四川、云南 2 个山区省^[1-4]。但在这些区域内疫情的分布也不均匀,探测出其中高风险区域已成为防治血吸虫病传播的重要任务之一。血吸虫传播过程的各个环节和因子均与其孳生的地理环境有着直接或间接的关系,而地理现象之间大部分都不具有独立性,这种空间分布特征的复杂性使得传统的数理统计方法无法很好地解决空间数据的关系问题,这就需要空间统计的分析方法^[5]。本研究采用 ArcGIS 软件中的空间统计分析模块和 SaTScan 软件综合探测 2008 年湖区 5 省血吸虫病传播的聚集区域,为防治提供依据,同时也为相关疾病聚集区域探测方法的建立提供参考。

资料与方法

1. 疫情数据:根据 2008 年全国血吸虫病疫情统计数据库中的“现有病例数”和“流行县人口数(万人)”两个变量,计算出江苏、安徽、江西、湖北、湖南 5 省各县/区血吸虫病感染率,并为每个县/区建立唯一的编号,以此作为空间分析的属性数据库。

2. 空间数据库:从 1:25 万全国县界电子地图中提取出上述 5 省的空间数据库,同时以属性数据库中的编号为参照,为该图层中的每个区县建立相同的惟一编号。在 ArcGIS9 (ESRI)中以相同的编号为匹配变量将属性数据库与空间数据库进行匹配,从而生成完整的空间分析数据库。

3. 空间聚集区域探测方法:空间自相关是指同一个变量在不同空间位置上的相关性,是空间单元属性值聚集程度的一种度量^[6]。Queiroz 等^[7]运用空间自相关对麻风病的高风险区域进行探测,取得较好结果。空间相关性通常采用全局和局部两种指标^[8]。全局指标主要用来探测整个研究区域的空间聚集模式,以单一的值来反映该区域的自相关程度;局部指标计算每一个空间单元与邻近单元同一属性的相关程度,检验研究区域内每个单元相对于整体范围而言其空间自相关是否足够显著,从而指出研究属性在空间高高、高低、低低的分布^[5,9,10]。由于全局指标有时会掩盖局部状态的不稳定性^[5],因此本研究采用 ArcGIS9 软件同时计算全局指标和局部指标,两个指标联合使用以探讨各单元之间的关系。SaTScan 是 Kulldorff 研究组开发的一款免费软件,它不仅分析疾病在空间或时空上的聚集分布并检验其分布是否具有统计学意义,而且还可以探测出的聚集的位置与范围^[11-14]。

(1) 全局空间自相关:本研究通过计算 Moran's I 反映全局空间自相关^[4]。公式:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{X})(x_j - \bar{X})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{X})^2} \quad E[I] = -\frac{1}{n-1}$$

$$V[I] = E[I^2] - E[I]^2 \quad z_i = \frac{I - E[I]}{\sqrt{V[I]}}$$

式中, n 为样本量, x_i 和 x_j 分别代表 i, j 点的属性值, w_{ij} 代表衡量空间事物关系的权重。但究竟采用哪种方法生成空间权重矩阵,至目前仍无根本的解决办法^[15]。本研究的权重矩阵选用反距离加权法,代表距离越远,相关性越低。以相邻属性权重的合计进行标化,并采用马氏距离法计算两点之间的距离。 $E[I]$ 为 I 的期望值, $V[I]$ 为 I 的理论方差^[16],通过 z 检验以判断其是否具有统计学意义^[17]。Moran's I 取值范围介于 $-1 \sim +1$ 之间, Moran's I 越接近于 1 则代表整体呈聚集性分布,样本之间关系越密切,性质越相似(高值聚集或低值聚集); Moran's I 越接近于 -1 则代表整体呈离散型分布,样本之间差异越大或分布越不集中^[9]。

(2) 局部空间自相关:本研究计算 Getis-Ord Gi 来反映局部空间自相关。局部 Getis-Ord Gi 是一种基于距离权重矩阵的局部空间自相关指标,能探测高值聚集和低值聚集^[18]。公式:

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\left[\frac{n \sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij} \right)^2}{n-1} \right]}}$$

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

式中, x_i 和 x_j 分别代表 i, j 点的属性值, w_{ij} 代表空间权重。本研究采用共同边界作为权重判断标准。在该方法中 Getis-Ord Gi 近似等于 z 值,可进行统计学 z 检验^[19]。当 $z > 0$ 且差异有统计学意义时,表示高值聚集;当 $z < 0$ 且差异有统计学意义时,表示低值聚集。

4. SaTScan 统计:SaTScan 软件常被用于空间疾病监测,是通过计算不同窗口下的对数似然比 (LLR),判断聚集区域的中心位置和范围。LLR 越大越有可能是聚集区域^[13,20]。公式:

$$L_{i,r} = \left(\frac{n_{i,r}}{\mu_{i,r}} \right)^{n_{i,r}} \left(\frac{N - n_{i,r}}{N - \mu_{i,r}} \right)^{N - n_{i,r}} \quad P = \frac{k+1}{M+1}$$

式中, $n_{i,r}$ 是扫描窗口的病例数, $\mu_{i,r}$ 是扫描窗口的期望病例数, N 是总病例数, M 为 Monte-Carlo 迭代的

次数, k 为 M 次迭代中扫描统计量大于或等于以真实资料为基础的扫描统计量的次数。本研究基于 SaTScan 8.0 软件的 Poisson 分布模型, 采用回顾性空间分析方法对 2008 年湖区 5 省的血吸虫病感染率进行分析, 探测出高风险区域的具体位置和范围。

结 果

1. 基本情况: 2008 年湖区 5 省血吸虫病流行县约有 261 个, 流行县总人口数约达 1.4 亿, 其中血吸虫病感染率约在 0 ~ 4% 之间(表 1、图 1)。

表 1 2008 年我国湖区 5 省血吸虫病基本情况

省份	流行县个数	流行县人口数($\times 10^7$)	现有病例数
江苏	71	3976.38	2 396
安徽	50	2157.37	36 812
江西	39	1796.99	90 322
湖北	63	3707.30	178 776
湖南	38	1892.72	92 932

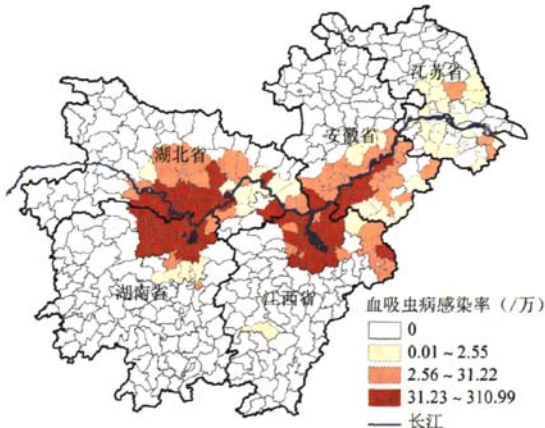


图 1 2008 年我国湖区 5 省血吸虫病感染率分布

2. 空间聚集区域分析:

(1) 全局空间自相关: 结果显示, Moran's I 为 0.47, $z=14.74$, $P<0.05$ 。从整体研究区域来看, 按照 $\alpha=0.05$ 的检验水准, 2008 年湖区 5 省的血吸虫病感染率分布呈现空间聚集性。

(2) 局部空间自相关: 按照 $\alpha=0.05$ 的检验水准, 2008 年湖区 5 省的血吸虫病感染率中有 50 个县(区)的局部空间自相关结果有统计学意义 ($P<0.05$)。其中湖北省聚集县(区)最多(17 个), 其次为江西省(15 个)、湖南省(14 个)、安徽省(4 个), 江苏省无聚集县(区)。50 个县(区)均 z 值 >0 , 提示均为高值聚集(图 2)。

3. SaTScan 统计: 共探测出 5 个聚集区域, 包括 39 个县(区)。其中有 34 个县(区)与局部空间自相关的结果吻合(图 3)。前三个聚集区域的相对危险度 (RR) 均 >3 , 最大对数似然比 (LLR) >1000 ($P<$

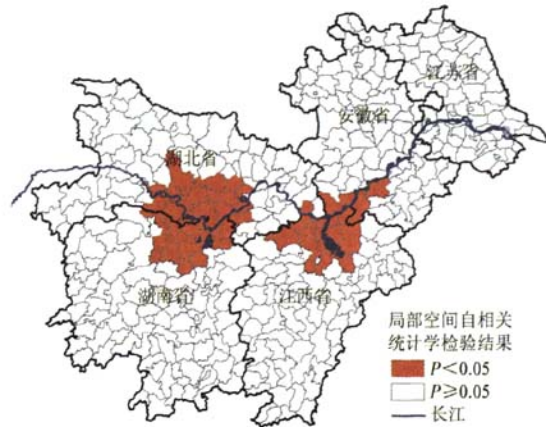


图 2 2008 年我国湖区 5 省血吸虫病感染率局部空间自相关结果

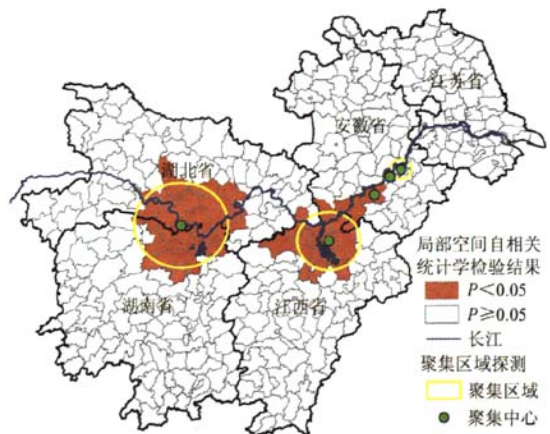


图 3 2008 年我国湖区 5 省血吸虫病 SaTScan 聚集区域探测结果与局部空间自相关结果合成图

0.05), 见表 2。第一个聚集区域位于长江与湖北、湖南省交界段, 此处聚集范围最大, RR 和 LLR 值最高。从此处开始向长江下游方向聚集范围逐渐缩小, RR 值也逐渐降低。

表 2 SaTScan 探测出我国湖区 5 省血吸虫病 5 个聚集区域

序号	聚集中心	半径 (km)	聚集范围 包括县(区)数	LLR	RR 值	P 值
1	湖北石首	116.951	23	195 949	8.19	0.001
2	江西都昌	79.613	12	46 867	4.53	0.001
3	安徽铜陵	-	1	1 080	3.03	0.001
4	安徽贵池	-	1	883	2.15	0.001
5	安徽繁昌	27.220	2	40	1.19	0.001

讨 论

了解血吸虫病的空间分布对于该病的控制具有重要意义^[21]。血吸虫病之所以在一个区域存在并引起流行, 与该地自然、人文地理条件适宜血吸虫病原体 and 寄生宿主的生长发育有密切关系^[22]。由于地理

因素之间存在的非独立性,采用传统的数理统计方法不能分析其空间数据的关系。疾病的空间聚集性分析是空间流行病学研究内容之一^[23],但目前也存在很多问题^[24],如各种聚集性方法以及权重指标的选择仍无统一标准^[15]。为此本研究对 2008 年湖区 5 省基于县(区)的血吸虫病疫情数据进行空间聚集性综合分析,并探测出具体聚集位置和范围。

本研究第一步采用全局空间自相关从研究整体上判断 2008 年湖区 5 省血吸虫病疫情具有空间聚集性;第二步采用局部空间自相关计算每个单位与单位之间的关系,得出有 50 个县(区)可能存在高值聚集;最后采用 SaTScan 软件探测出聚集区域的中心以及聚集范围,并与局部空间自相关的结果进行比较。发现两种方法探测的聚集区域较为吻合。5 个主要聚集中心分布于长江流经湖北、湖南、江西、安徽 4 省的区域,聚集范围从上游到下游呈现逐渐缩小的趋势,最大的聚集区域位于湖北省与湖南省交界地带。但是局部自相关结果与 SaTScan 探测的结果略有不同。究其原因这是由于两种方法探测聚集区域的原理各异所致。局部空间自相关的是通过计算每一个单元格与邻近单元格就某一属性的相关程度,从而判断其空间模式。而 SaTScan 探测是以 LLR 为基础,运用 Monte-Carlo 迭代模拟进行统计学检验。全局空间自相关可以从整体上把握空间数据是否具有空间聚集性;局部空间自相关可以探测出各个单元格之间的关系,从而判断其具体聚集区域;SaTScan 可以计算出其具体的中心位置和范围。三种方法联合使用,逐步深入,使得结果呈现更加全面、系统。

综合以上结果提示,我国湖区 5 省沿江地带仍然是血吸虫病空间聚集的主要区域,尤以湖北、湖南省交界地带空间聚集性最高、范围最大。该区域的自然环境和社会行为等相关因素可能更利于血吸虫病的传播,人群感染血吸虫病的风险也较大。因此,应重点针对这些聚集区域,加强血吸虫病及其相关社会、环境等因素的监测,实施有力措施阻断血吸虫病的传播。在研究方法上,联合使用全局空间自相关、局部空间自相关以及 SaTScan 统计软件可以更加系统、全面的探测血吸虫病的空间聚集区域。

参 考 文 献

[1] Wang LD, Chen HG, Guo JG, et al. A strategy to control transmission of *Schistosoma japonicum* in China. *N Engl J Med*, 2009, 360(2): 121-128.
 [2] Wu XH, Xu J, Zheng J, et al. Challenges and control strategies in areas of schistosomiasis transmission controled and interrupted in China. *Chin J Schistosomiasis Control*, 2004, 16(1): 1-3. (in Chinese)
 吴晓华, 许静, 郑江, 等. 中国血吸虫病传播控制与阻断地区面临的挑战及对策. *中国血吸虫病防治杂志*, 2004, 16(1): 1-3.

[3] Zheng J. Challenges and prospects of control strategies in areas of schistosomiasis in China. *J Trop Dis Parasitol*, 2004, 2(4): 193-197. (in Chinese)
 郑江. 我国血吸虫病防治策略面临的挑战及发展方向. *热带病与寄生虫学*, 2004, 2(4): 193-197.
 [4] Chen Z, Zhou XN, Yang K, et al. Strategy formulation for *Schistosomiasis japonica* control in different environmental settings supported by spatial analysis: a case study from China. *Geospat Health*, 2007, 1(2): 223-231.
 [5] Qi XP, Zhou MG, Hu YS, et al. Spatial hotspot exploration on digestive tract cancer mortality with geographic information system. *Geogr Res*, 2010, 29(1): 181-187. (in Chinese)
 戚晓鹏, 周脉耕, 胡以松, 等. 应用地理信息系统探测消化道癌症死亡率空间聚集性. *地理研究*, 2010, 29(1): 181-187.
 [6] Getis A, Ord K. The analysis of spatial association by use of distance statistics. *Geogr Analysis*, 1992, 24(3): 189-206.
 [7] Queiroz JW, Dias GH, Nobre ML, et al. Geographic information systems and applied spatial statistics are efficient tools to study Hansen's disease (Leprosy) and to determine areas of greater risk of disease. *Am J Trop Med Hyg*, 2010, 82(2): 306-314.
 [8] Zhang SL, Zhang K. Contrast study on Moran and Getis-Ord indexes of local spatial autocorrelation indices. *J Geodesy Geodynamics*, 2007, 27(3): 31-34. (in Chinese)
 张松林, 张昆. 空间自相关局部指标 Moran 指数和 G 系数研究. *大地测量与地球动力学*, 2007, 27(3): 31-34.
 [9] Zhang SL, Zhang K. Comparison between general Moran's index and Getis-Ord general G of spatial autocorrelation. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2007, 46(4): 93-97. (in Chinese)
 张松林, 张昆. 全局空间自相关 Moran 指数和 G 系数对比研究. *中山大学学报(自然科学版)*, 2007, 46(4): 93-97.
 [10] Zhang SL, Zhang K. Contrast study on the local indices of spatial autocorrelation. *Stat Res*, 2007, 24(7): 65-67. (in Chinese)
 张松林, 张昆. 局部空间自相关指标对比研究. *统计研究*, 2007, 24(7): 65-67.
 [11] Kulldorff M, SaTScan TM. User guide for version 8.0 [DB/CD]. 2009.
 [12] Coleman M, Coleman M, Mabuza AM, et al. Using the SaTScan method to detect local malaria clusters for guiding malaria control programmes. *Malar J*, 2009, 8: 68-73.
 [13] Kulldorff M. A spatial scan statistic. *Communications in Statistics-Theory and Methods*, 1997, 26(6): 1481-1496.
 [14] Zhang ZJ, Peng WX, Zhou YB, et al. Study on the quantified indices to describe the distributional status of diseases in the spatial point pattern analysis. *Chin J Prev Med*, 2008, 42(6): 422-426. (in Chinese)
 张志杰, 彭文祥, 周艺彪, 等. 基于空间点模式分析的疾病分布状态的量化指标研究. *中华预防医学杂志*, 2008, 42(6): 422-426.
 [15] Chen YG. Reconstructing the mathematical process of spatial autocorrelation based on Moran's statistics. *Geogr Res*, 2009, 28(6): 1449-1463. (in Chinese)
 陈彦光. 基于 Moran 统计量的空间自相关理论发展和方法改进. *地理研究*, 2009, 28(6): 1449-1463.
 [16] Bivand R. A review of spatial statistical techniques for location studies. *Citeseer*, 1998.
 [17] Ward MP, Carpenter TE. Techniques for analysis of disease clustering in space and in time in veterinary epidemiology. *Prev Vet Med*, 2000, 45(3-4): 257-284.
 [18] Ord JK, Getis A. Testing for local spatial autocorrelation in the presence of global autocorrelation. *J Reg Sci*, 2001, 41(3): 411-432.
 [19] Tsai PJ, Lin ML, Chu CM, et al. Spatial autocorrelation analysis of health care hotspots in Taiwan in 2006. *BMC Public Health*, 2009, 9: 464-476.
 [20] Abrams AM, Kleinman KP. A SaTScan macro accessory for cartography (SMAC) package implemented with SAS software. *Int J Health Geogr*, 2007, 6(6): 1-8.
 [21] Zhang ZJ. Spatial analysis on the high risk regions for *Schistosomiasis japonica* and identification of active transmission sites. Fudan University, 2008. (in Chinese)
 张志杰. 湖沼地区血吸虫病高风险区域的空间分析及重点钉螺孳生地的探测. 复旦大学, 2008.
 [22] Zhao A, Jiang MX, Jian MF, et al. Progress and prospects of schistosomiasis-medical geography. *Prog Geogr*, 2010, 29(1): 45-51. (in Chinese)
 赵安, 蒋梅鑫, 简敏非, 等. 血吸虫病医学地理研究的回顾与展望. *地理科学进展*, 2010, 29(1): 45-51.
 [23] Yang K, Wang XH, Wu XH, et al. Application of spatial epidemiology techniques in the research and prevention of schistosomiasis. *Chin J Public Health*, 2007, 23(8): 1017-1021. (in Chinese)
 杨坤, 王显红, 吴晓华, 等. 空间流行病学技术在血吸虫病防治研究中的应用. *中国公共卫生*, 2007, 23(8): 1017-1021.
 [24] Elliott P, Wartenberg D. Spatial epidemiology: current approaches and future challenges. *Environ Health Perspect*, 2004, 112(9): 998-1006. (收稿日期: 2010-04-08)
 (本文编辑: 张林东)