· 专题报道之六 ·

中国 2002—2011 年肺结核流行状况 GIS 空间分析

山珂,徐凌忠,盖若琰,王兴洲,于小龙,祁华金,崔伊萌,卢怡帆,徐融飞

摘　要:目的　分析 2002—2011 年中国肺结核发病的流行状况特征。**方法**　运用 ArcGis 10.0 软件探索性数据分析和空间分析方法,对中国肺结核疫情数据进行空间显示和统计分析。**结果**　肺结核发病图和空间自相关分析结果显示,2002—2011 年肺结核发病存在高发区和低发区,并存在空间聚集性;局部 Local Moran′s *I* 自相关分析结果表明,新疆、贵州、广西地区为高高值聚集,北京、天津、河北、山东、辽宁、江苏为低低值聚集;局部热点分析 Getis-Qrd General G_i^* 结果表明,全国肺结核疫情存在 13 个"热点"区域,其中 4 个"正热点"区域分别为新疆、西藏、云南、广西,9 个"负热点"区域分别为北京、天津、河北、辽宁、山东、山西、河南、江苏、内蒙古。**结论**　中国肺结核发病具有明显的地域分布,存在中度以上的空间聚集性,应针对其地域特点进一步研究。

关键词:肺结核;地理信息系统(GIS);空间自相关分析;趋势分析
中图分类号:R 521　**文献标志码:**A　**文章编号:**1001-0580(2014)04-0388-04　DOI:10.11847/zgggws2014-30-04-02

Spatial analysis on tuberculosis epidemic in China based on geographic information system, 2002 – 2011

SHAN Ke, XU Ling-zhong, GAI Ruo-yan, et al (*Department of Social Medicine and Health Managemet, Branch of Health Management and Woman and Children Health Care, School of Public Health, Shandong University, Ji′nan, Shandong Province*, 250012, China)

Abstract: Objective　To analyze the epidemic characteristics of tuberculosis epidemic in China in 2002 – 2011. **Methods**　The data on tuberculosis incidence in China were collected and linked to the digital map at administrative county level with ArcGIS 10.0 software for spatial analysis. **Results**　The incidence maps of tuberculosis and spatial autocorrelation analysis showed that the spatial aggregation existed in tuberculosis epidemic during 2002 – 2011. Local Moran′s I and Getis-Ord General G spatial autocorrelation analysis identified the "high incidence region" and "hot pot region". **Conclusion**　The incidence of tuberculosis is of obviously geographical distribution in China. And spatial clustering nature of tuberculosis incidence between 2002 and 2011 in China appeared to be moderate or high. However, further investigation is needed to address this changing trend.

Key words: pulmanary tuberculosis; geographic information system; spatial autocorrelation analysis; trend analysis

结核病是一种严重危害人类健康的慢性传染病,素有"白色瘟疫"之称。中国是全球 22 个结核病高负担国家之一,结核病人数位列世界第二位[1]。目前,国内外分析肺结核流行特点和暴发趋势多采用一般性描述和时间趋势分析,缺乏定量分析其空间特性和聚集特点[2],而地理信息系统(geographical information system, GIS)技术是将有关空间数据管理和空间信息分析相结合的一种计算机技术[3]。本研究将地理空间信息与我国结核病发病属性相结合,描述肺结核疫情的空间分布并进行相关空间分析,为识别和预测肺结核重点防控区域提供理论依据。

1　资料与方法

1.1　*资料来源*　收集 2002—2011 年中国大陆地区(不包括香港、澳门、台湾)肺结核发病情况资料及其相关人口资料,数据均来自各 2003—2012 年《中国统计年鉴》、《中国卫生统计年鉴》以及《我国卫生事业发展统计公报》。以全国 1∶100 万省界矢量地图作为基础地图。利用 ArcGIS 10.0 软件将整理完善的 2002—2011 年中国肺结核发病资料与基础地图上的行政区划进行匹配,建立相应的空间数据库和属性数据库。

1.2　*方法*　应用 ArcGIS 10.0 软件对肺结核数据进行二次多项式拟合插值,得到全国肺结核发病的三维立体趋势图。根据我国行政区划地理相邻关系,利用基于反距离的空间概念化方法生成空间权重矩阵。在定义好的权重矩阵基础上,分别计算每年全局 Moran′s *I* 自相关指数和全局 General G 指数,并进行假设检验,判断省级水平上肺结核疫情是否存在聚集性。同时在权重矩阵基础上,进行局部 Local Moran′s *I* 自相关分析和局部热点分析,绘制肺结核疫情聚集分布图和热点分布图。

1.3　*统计分析*　(1)采用 Moran′s *I* 空间自相关分

作者单位:山东大学公共卫生学院卫生事业管理与妇幼卫生学系社会医学与卫生管理科室,山东 济南 250012
作者简介:山珂(1988 –),女,山东日照人,硕士在读,研究方向:社会医学与卫生管理。
通讯作者:徐凌忠,E-mail:lzxu@sdu.edu.cn
数字出版日期:2013 – 7 – 17 9:38
数字出版网址:http://www.cnki.net/kcms/detail/21.1234.R.20130717.938.009.html

析方法[4]。空间自相关分析分为全局空间自相关分析与局部空间自相关分析。全局空间自相关是对某一现象在整个研究区域空间特征的描述;局部空间自相关是研究范围内各空间位置与各自周围邻近位置的同一现象的相关性。Moran 指数的取值范围近似为 -1 ~1, <0 表示负相关,0 表示不相关, >0 表示正相关。越接近 -1 则代表单元间的差异越大或分布越不集中,越接近 1 则代表单元间的关系越密切,性质越相似,接近 0 则代表单元间不相关。对于全局 Moran 指数,可以用标准化统计量 $Z(I)$ 来检验空间自相关的显著性水平。(2)热点分析 Getis-Ord G。由于 Moran 指数不能判断空间数据是高值聚集还是低值聚集,Getis 和 Ord 于 1992 提出了全局 G 系数。G 系数一般采用距离权重,要求空间单元的属性值为正。类似于全局 Moran 指数,也可对 G 系数进行标准化正的 $Z(G)$ 表示存在高值聚集,负的 $Z(G)$ 表示存在低值聚集[4]。

2　结　果

2.1　中国结核病疫情空间分布(图 1)　在 ArcGIS 桌面窗口 ArcMap 中使用自然分割法将各省肺结核平均发病率均分成 5 个相同等级,并以由浅到深的颜色表示发病率从低到高,同时使用柱形图表示各个年份各省肺结核发病率的变化趋势。图中提示肺结核发病存在高发区和低发区,高发区域可能有潜在的危险因素。整体趋势上病情大致呈现由南向北逐渐升高,由西向东先下降后上升的趋势。新疆、贵州、海南、广西、重庆、西藏、甘肃、四川、湖北、广东 10 省(直辖市/自治区)发病指数较高,宁夏、江苏、河北、辽宁、北京、山东、上海、天津 8 省(直辖市/自治区)发病指数较低。表明高发病率地带主要集中在西南、西北地区,东部沿海地区较低,这种分布模式在一定程度上反映了空间聚集的存在。

2.2　中国肺结核疫情趋势分析　以肺结核疫情为因变量 z,地理位置的横坐标(经度)和纵坐标(纬度)分别为自变量 x、y 进行趋势分析。结果表明,2002—2011 年肺结核疫情总体分布自南向北呈先下降后上升的趋势,自西向东逐渐降低的趋势。西南地区和西北地区发病率情况较高,表明存在区域聚集性。

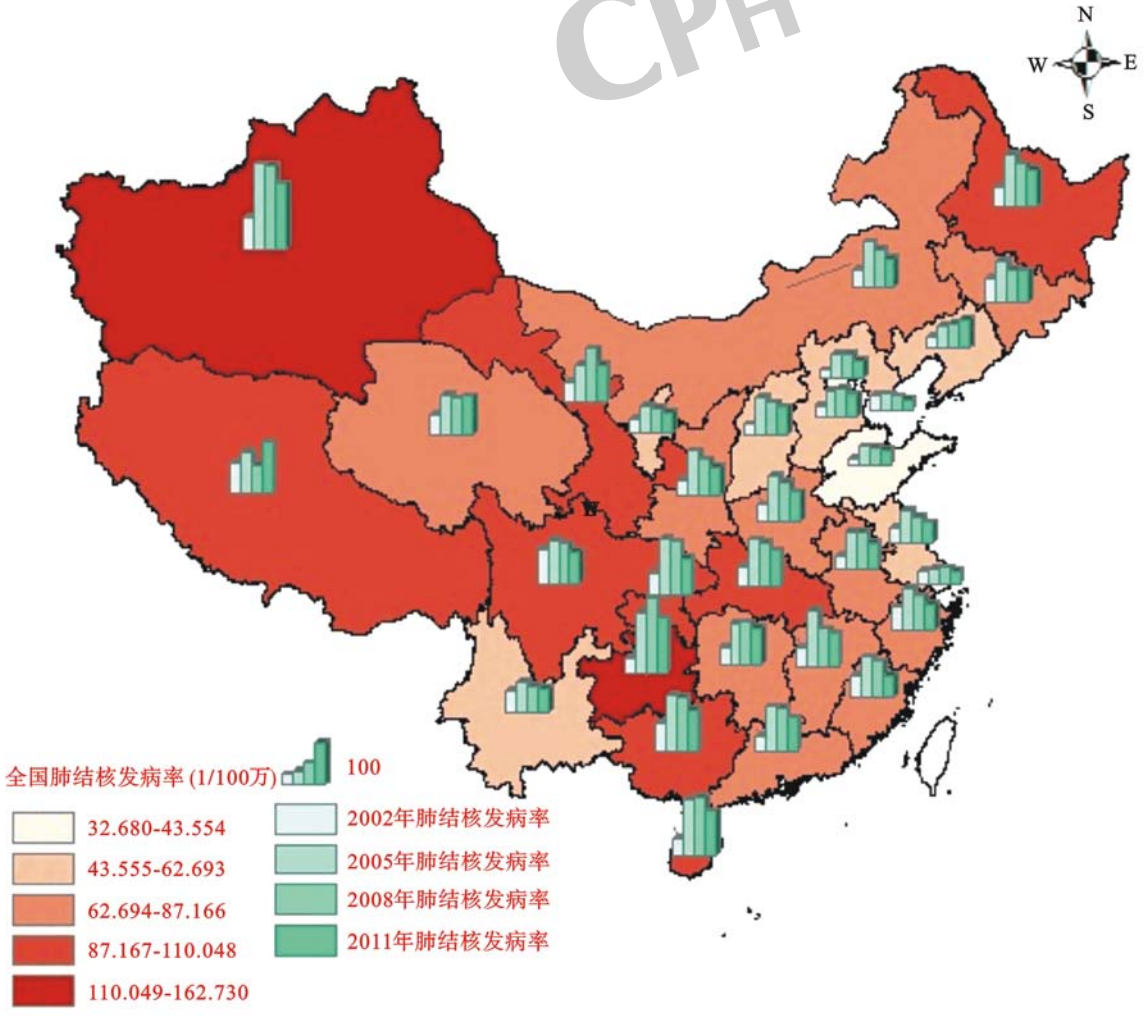

图 1　中国 2002—2011 年肺结核发病分布及变化趋势

2.3　空间自相关分析

2.3.1　全局 Moran′s *I* 和局部 Local Moran′s *I* 自相关分析(表 1、图 2)　全局 Moran′s *I* 自相关分析结果显示(表 1),2002—2011 年肺结核空间自相关指数分别为 0.409 862、0.356 617、0.352 067、0.351 000、0.339 927、0.351 737、0.298 502、0.332 077、0.360 194、0.368 683,经蒙特卡洛检验,*P* 值均 <0.05,表明肺结核疫情在区域范围内自相关关系有统计学意义;Moran′s *I* 指数为 0.3 ~0.7,表现为中等程度及以上的空间聚集性。局部 Local Moran′s *I* 自相关分析结果显示(图 2),全国肺结核存在空间聚集,其中新疆、贵州、广西地区为高高值聚集,北京、天津、河北、山东、辽宁、江苏为低低值聚集。

表 1　2002—2011 年中国肺结核疫情全局 Moran′s *I* 相关分析及变化趋势

年份	Moran *I* 指数	Var(*I*)	*Z* 值	*P* 值
2011	0.368 683	0.004 968	5.703 491	0.000 000
2010	0.360 194	0.004 898	5.622 783	0.000 000
2009	0.332 077	0.004 819	5.263 968	0.000 000
2008	0.298 502	0.004 863	4.758 462	0.000 002
2007	0.351 737	0.004 902	5.499 796	0.000 000
2006	0.339 927	0.004 889	5.338 381	0.000 000
2005	0.351 000	0.004 921	5.478 689	0.000 000
2004	0.352 067	0.004 933	5.370 595	0.000 000
2003	0.356 617	0.005 061	5.481 185	0.000 000
2002	0.409 862	0.005 074	5.188 601	0.000 000

图 2　中国肺结核发病率 Local Moran′s *I* 相关分析

2.3.2　全局热点分析 Getis-Ord General G 和局部热点分析 Getis-Ord General Gi*(表 2、图 3)　全局热点分析 Getis-Ord General Gi 结果表明,全国肺结核发病率 General G 指数 = 0.000 001,在 2002—2004 年和 2006—2011 年均 *Z* <-1.96,均 *P* <0.05,表明肺结核疫情在区域范围内相关关系有统计学意义,存在“热点”区域。局部热点分析 Getis-Ord General Gi* 结果表明,全国肺结核疫情局部自相关存在 13 个“热点”区域;其中 4 个“正热点”区域分别为新疆、西藏、云南、广西,与周围区域疾病分布呈高高相邻,即疾病高发区;9 个“负热点”区域分别为北京、天津、河北、辽宁、山东、山西、河南、江苏、内蒙古,即疾病低发区。

表 2　2002—2011 年全国肺结核疫情空间自相关分析及变化趋势

年份	General G 指数	Var(*I*)	*Z* 值	*P* 值
2011	0.000 001	0.000 000	-2.952 991	0.009 515
2010	0.000 001	0.000 000	-2.651 389	0.008 016
2009	0.000 001	0.000 000	-2.415 671	0.015 706
2008	0.000 001	0.000 000	-1.570 517	0.116 295
2007	0.000 001	0.000 000	-1.751 757	0.079 816
2006	0.000 001	0.000 000	-1.896 166	0.057 938
2005	0.000 001	0.000 000	-1.833 596	0.066 714
2004	0.000 001	0.000 000	-2.300 084	0.040 452
2003	0.000 001	0.000 000	-2.732 239	0.006 291
2002	0.000 001	0.000 000	-2.896 438	0.003 774

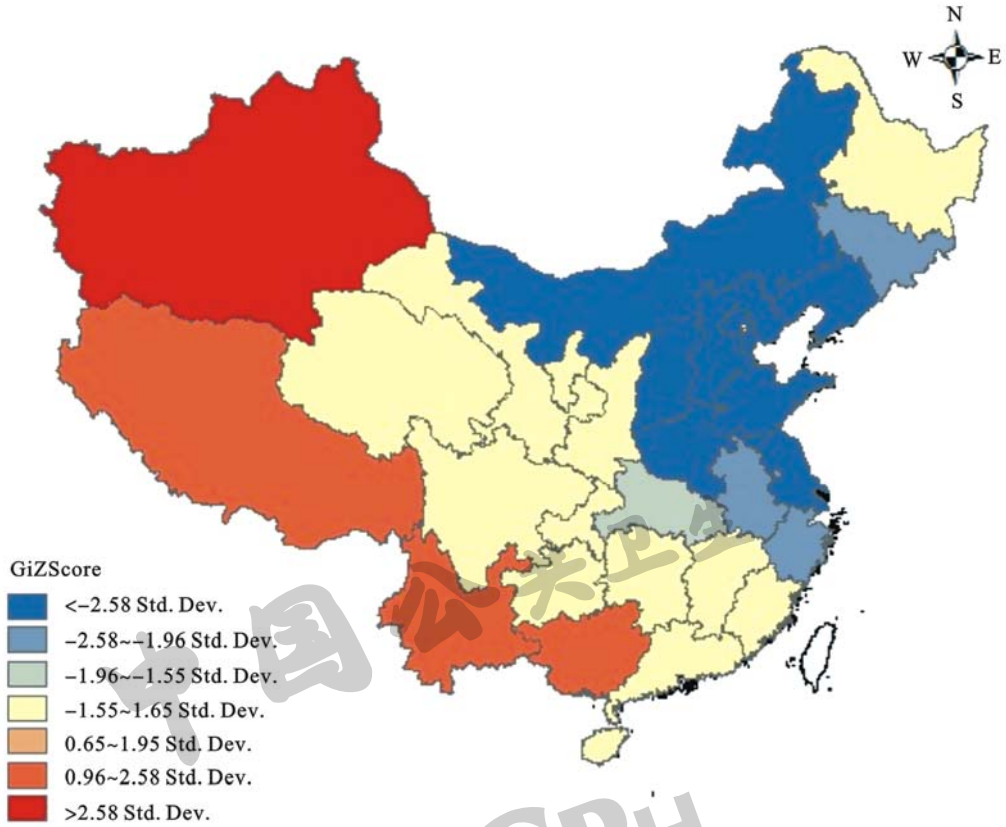

图 3　中国肺结核发病率局部热点分析

3　讨　论

本研究结果表明，2002—2011 年全国肺结核疫情在省级层面无论是从各时间维度（年份）分析，还是从 10 年总发病率分析，均存在空间聚集。对 2002—2011 年肺结核疫情趋势分析，得到肺结核疫情总体分布自南向北呈先下降后上升趋势，自西向东逐渐降低趋势。西南和西北地区发病率情况较高，这可能与不同区域相关部门的重视程度以及气候、环境因素相关[5]。Moran′s *I* 自相关分析表明，全国肺结核存在空间聚集，其中新疆、贵州、广西地区为高高值聚集，北京、天津、河北、山东、辽宁、江苏为低低值聚集。高高值聚集区域主要集中在西部地区，高发病率可能与经济状况以及医疗制度设施不健全相关；低低值聚集区域以北京为中心聚集分布，低发病率可能与医疗水平、文化素质以及经济状况相关。热点分析结果表明，存在 13 个“热点”区域，其中 4 个“正热点”区域分别为新疆、西藏、云南、广西，9 个“负热点”区域分别为北京、天津、河北、辽宁、山东、山西、河南、江苏、内蒙古，结果与Moran′s *I*相关分析结果基本一致，与文献[6]结果相似。本研究采用空间自相关方法对 2002—2011 年全国肺结核疫情进行空间聚集性分析，仅从省级层面对聚集性进行分析，揭示了肺结核疫情的空间流行状况和浅层原因，而对其形成原因和相关机制有待进一步探索[7-8]。

参考文献

[1] 应倩，陈坤. 应用 G 统计量分析 2006 年浙江省肺结核的空间异质性[J]. 中华预防医学，2012，46(6)：524－528.
[2] 于小龙，郭振，郑文贵，等. 山东省结核病患者服药依从性相关因素分析[J]. 中华预防医学，2012，46(7)：661－662.
[3] Prasannakumar V，Vijith H，Charutha R，et aL. Spatio-temporal clustering of road accidents：GIS based analysis and assessment [J]. Procedia-Social and Behavioral Science，2011，21 (7)：317－325.
[4] 周晓农. 空间流行病学[M]. 北京：科学出版社，2009：237－239.
[5] 唐成艳，黄天壬，朱小东，等. 应用空间自相关分析研究广西壮族自治区肝癌的空间异质性分布特征[J]. 中华流行病学杂志，2009，30(2)：167－170.
[6] 成月佳，鲍昌俊，梁祈，等. 江苏省病毒性肝炎流行趋势 GIS 空间分析[J]. 中国公共卫生，2011，27(7)：873－876.
[7] 张文增，李长青，冀国强，等. 空间扫描统计量在手足口病空间聚集性研究中的应用[J]. 中国卫生统计，2012，29(4)：507－513.
[8] 黄开勇，唐咸艳，王晓敏，等. 桂林市道路交通伤害 GIS 反距离加权插值法分析[J]. 中国公共卫生，2012，28(5)：576－578.

收稿日期：2013－02－20　（潘雯编辑　郑新校对）