

论 文

北京市 SARS 流行的特征与时空传播规律

曹志冬^①, 曾大军^{①*}, 郑晓龙^①, 王全意^②, 王飞跃^①, 王劲峰^③, 王小莉^②^① 中国科学院自动化研究所复杂系统与智能科学重点实验室, 北京 100190;^② 北京市疾病预防控制中心, 北京 100013;^③ 中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室, 北京 100101

* 联系人, E-mail: zengdaniel@gmail.com

收稿日期: 2009-05-19; 接受日期: 2009-10-28

美国自然科学基金(编号: IIS-0839990, IIS-0428241)、美国国土安全部项目(编号: 2008-ST-061-BS0002)、国家科技重大专项(编号: 2009ZX10004-315, 2008ZX10005-013)、中国科学院项目(编号: 2F07C01, 2F08N03)、博士面上项目(编号: 20080440559)、国家高技术研究发展计划(编号: 2006AA010106)与国家自然科学基金(编号: 60621001, 40901219)资助

摘要 传染病流行的时空传播规律反映了人与传染病斗争博弈的动态过程, 深刻认识时空传播规律对于传染病的科学防控与应急管理有重要意义. 对 2003 年北京市的流行病学调查数据进行数字化得到了 2321 例 SARS 感染者的时空数据, 重现了 SARS 在北京市流行的时空传播过程, 得到了精确到街道办、乡和镇的 SARS 发病率图, 并对结果进行了 Bayesian 调整和空间平滑, 使之更好的反映出 SARS 传染病在地理空间上的风险分布与趋势; 综合运用多种空间分析方法研究了 SARS 传播的空间模式与时空传播规律, 并结合应急防控措施阐述了时空传播背后的驱动因素, 评价了不同时期的应急策略对疫情控制的效果. 结果表明: (1) SARS 传播在西北-东南方向(偏南北向)的扩散速度明显弱于东北-西南方向; (2) SARS 传播风险在空间上存在显著正相关, 且经历了弱-强-弱的变化过程; (3) 在空间上形成了两个高风险传播的热点区域——城市中心地带与东部城郊地带, 两个风险区域的形成与演化特征各不相同但相互影响; (4) 错过了在 3 月份将 SARS 流行消灭在萌芽状态的最佳时机, 然而, 疫情中后期的应急策略非常有效, 不仅阻止了疫情的增长态势, 而且有效遏止了其在空间上的扩散.

关键词

SARS

北京

发病率

空间分析

时空传播规律

防控措施

严重急性呼吸系统综合症(Severe Acute Respiratory Syndrome, SARS)是一种冠状病毒感染的新型传染性疾病, 其主要传播途径是近距离的空气飞沫接触、患者的呼吸道分泌物和密切接触, 尤其是密闭环境下的近距离接触极易导致感染. SARS 的传播流行是一系列复杂的自然因素与社会因素综合作用的结果, 它在时空上的传播扩散会呈现出一定规律, 这既反映了正常人群经传播途径受到感染的时序变化

过程, 又体现了正常人群与感染人群在空间上的分布与交互活动导致的空间扩散过程. 定量研究 SARS 流行的空间模式与时空传播规律不仅可以更加全面深刻的认识 SARS 流行的规律与内在传播机制, 评估不同地区的正常人群受 SARS 感染的风险等级, 而且能用于分析 SARS 流行过程中的各类应急防控措施对疫情传播的影响, 评比政策得失, 为未来可能再次出现的 SARS 流行及其他新发突发急性传染病流行

引用格式: Cao Z D, Zeng D J, Zheng X L, et al. Spatio-temporal evolution of Beijing 2003 SARS epidemic. Sci China Earth Sci, 2010, doi: 10.1007/s11430-010-0043-x

的应急决策提供科学依据。

2003 年突发的 SARS 流行是新中国成立后严重的一次公共卫生安全危机,也是进入 21 世纪以来人类面临的影响较为广泛的一次传染病全球性流行事件。五年来,世界各地的大量科学家分别从病理学^[1]、分子生物学^[2]、临床医学^[3]、流行病学^[4]、卫生经济学^[5]及应急管理^[6]等方面展来了对 SARS 的全面研究,取得了许多重要成果,然而,仍有许多未知谜团有待深入发掘。

中国是受 SARS 危害最大的国家,一半以上的 SARS 感染者出现在中国大陆,其中,SARS 在北京市的流行最为严重,北京市的 SARS 感染者占全球感染总数约 30%,在特异性疫苗研制成功并得以普及以前,揭示北京市 SARS 流行的时空传播规律、阐述其内在传播机制及定量评价应急政策得失具有重要意义,尤其是面临着禽流感、手足口病、甲型 H1N1 等与 SARS 类似的新发突发传染病的严重威胁。北京市是一座人口达千万以上的特大型城市,经济全球化背景下,北京市面临着更为严峻的传染病暴发流行的威胁。因此,以 2003 年北京市暴发的 SARS 流行事件作为样本来研究其时空传播规律不仅意义重大,而且具有很强的示范性。

时至今日,虽有大量针对北京 SARS 流行规律的研究成果,但大多基于传统流行病学调查的经典统计分析^[7-9]、时序过程的数学建模^[10-13]、仿真模拟计算^[14]与预测^[15,16]、影响 SARS 流行的风险因子分析^[17,18]及 SARS 流行对国民经济影响^[19,20]等,对北京市 SARS 流行的空间模式特征与时空传播规律的研究较少。已有研究中,有部分学者展开过针对 SARS 流行规律的时空分析研究,如针对北京市的上万名 SARS 密切接触者的多维空间分布、聚类及相关性特征分析^[21-24],及对北京市在城区与城郊的 SARS 流行特征差异的对比及其成因探讨^[25]。然而,这些研究尚存在一些不足之处,具体为:① 密切接触者的时空规律能够一定程度上反映 SARS 疫情的时空扩散特征,但不能完全代替针对 SARS 感染者本身的研究,后者能够更全面更清晰的反映出 SARS 疫情的时空传播规律;② 以城区和城郊来对 SARS 流行的时空特征进行差异化分析能够得到很多有用信息,然而,空间统计单元(大的行政分区,如朝阳区、海淀区等)较为粗糙,致使信息粒度过粗,无法揭示街道办、乡和镇等层级的时空传播规律,而这些信息对传染病

防控的指导意义更好;③ 已有研究大多是时、空分离的时空规律研究,即时序特征与空间分布特征结合的模式,没有将时空信息统一起来综合研究时空联动情况下的传播扩散规律;④ 缺少将 SARS 疫情的时空传播规律与政府的应急政策相结合的综合分析,以评比政策之得失。

基于以上分析,并结合与北京市疾病预防控制中心合作并数字化得到的 2321 例 SARS 感染者的时空数据,综合利用最新发展的地理信息技术(Geographic Information System, GIS)与现代空间统计分析方法(Spatial Statistical Analysis, SSA),全面探讨了北京市 SARS 疫情的疫情传播的空间变异模式、空间聚类特征及其异质性与时空传播规律,深刻论述了影响传播的驱动因素,并将 SARS 流行的时空传播规律及其流行病学解释与疫期中的应急防控策略相结合,评比了政策之得失。

1 数据与空间化处理

1.1 流行病学调查数据

实验数据来源于北京市疾病预防控制中心(CDC),主要根据 SARS 感染者(或亲属)的回顾性调查获取,共计 2444 例感染者数据。感染者发病的时间范围从 2003 年 3 月 8 日到 2003 年 5 月 28 日(图 1)。流行病学调查内容包括年龄、性别、家庭住址、工作住址和出现症状的时间等信息。由图 1 知, SARS 疫情经历了前期缓慢增长、中期迅速攀升与后期迅速下降直至完全消失的变化过程,在 4 月中下旬达到高峰。

1.2 时空数据

流行病学调查得到的原始数据中感染者的时间与空间信息表达为文本形式,无法用于定量分析

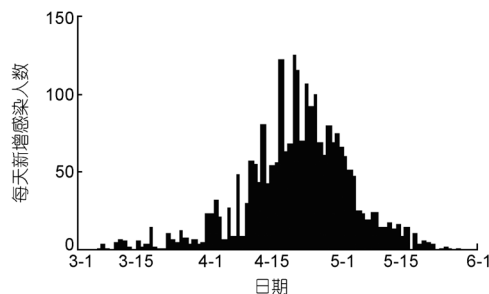


图 1 2003 年每天新增 SARS 感染者人数的时序变化

SARS 流行的时空传播规律, 因此, 需要采用空间化处理方法对 SARS 感染者的时间与空间信息进行时空数据的数字化处理. 家庭是人们生活最频繁的地方, 人的生活范围一般以家庭为圆心, 事实上, 家庭传播是 2003 年 SARS 传播的主要途径之一, 以**家庭地址**作为 SARS 暴发的**空间线索**来考查 SARS 流行的时空传播规律是合乎情理的. 另外, 原始数据中感染者的家庭住址信息最为充分完备, 其他空间信息(如工作单位地址)存在严重的语焉不详、数据缺失等不足. 因此, 本文采用感染者的**发病时间**作为 SARS 时空数据的时间信息, 采用感染者的家庭住址作为 SARS 时空数据的空间信息, 以此来构成 2003 年北京市 SARS 流行传播的**GIS 时空数据**. 由于流行病学调查数据采集过程中的各种不确定性因素, 使少数家庭地址信息出现缺失、太泛或错误, 为此, 论文采用表 1 中的空间化处理原则来指导时空数据的数字化.

根据表 1 的空间化处理原则, 以采用人工方法对

地名进行匹配, 得到了 2321 例 SARS 感染者在北京市的空间分布情况(图 2(a)), 空间化率约为 $2321/2444 \times 100\% = 95\%$, 北京市的 SARS 感染者大多集中在人口稠密的城市中心地带, 周边郊区及远郊区县的患者数量非常少. 为消除少数远离 SARS 传播中心的感染者对研究分析产生的不确定性影响, 本文选用以 SARS 感染者的**空间几何中心为圆心**、两倍距离标准差为半径的疫区中心地带的感染者为研究对象(图 2(b)), 范围涵盖北京市五环以内所有区域, 研究范围内的感染者数量为 2227 例, 占有可定位的时空数据总量的 96%, 占北京市所有感染者人数的 91.1%.

2 北京市 SARS 流行的发病率图

2.1 发病率图

发病率(Morbidity rate)是指一定时期内新病例的

表 1 SARS 感染者数据的空间化处理原则^{a)}

所在省、市	家庭住址	单位住址	空间化处理方法	结果统计	
				数量	比例
北京市	√	√	家庭地址	2158	88.30%
	√	×	家庭地址	0	0
	×	√	单位地址	163	6.67%
	×	×	弃之不用	0	0
非北京市	√/×	√/×	弃之不用	123	5.03%

a) √表示信息完整可靠, 可以进行空间定位; ×表示信息缺失、太泛或错误, 无法进行空间定位

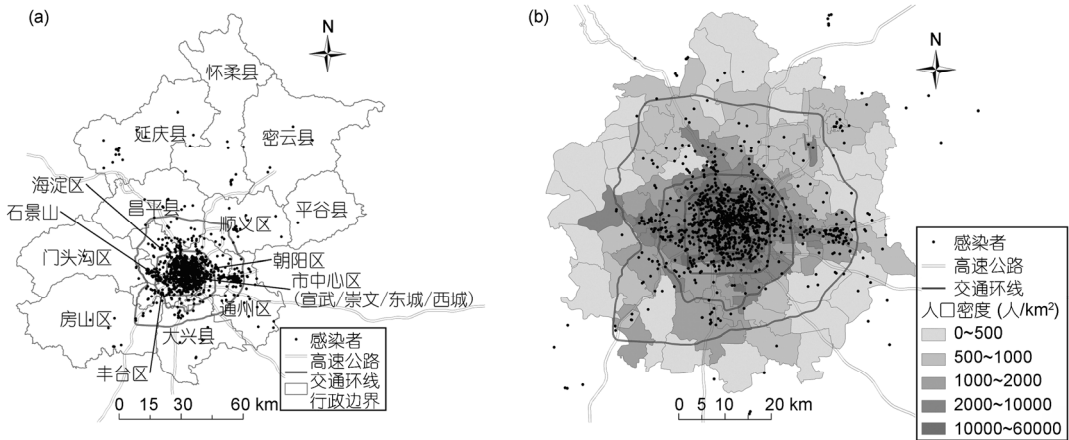


图 2 2003 年北京市 SARS 感染者的空间分布

(a) 北京市 14 区 2 县的分布, 共计 2231 例 SARS 感染者; (b) 疫情传播中心的分布, 共计 2227 例 SARS 感染者

发生频率, 计算公式为: 发病率=观察期间内发生的新病例数/同期平均人口数.

论文中的空间统计单元精确到街道办、乡和镇(图 2(b)), 最小统计单元的面积约为 1.03 km^2 , 最大面积约为 163 km^2 . 空间统计单元上的 SARS 发病人数根据 SARS 感染者的空间位置与空间统计单元所在的多边形进行叠加得到. 空间统计单元上的常住人口数采用中国第 5 次 1% 人口抽样调查时的人口统计数据. 直接将 SARS 流行期内空间统计单元上的感染人数除以所在统计单元内的常住人口数得到 2003 年北京市的 SARS 发病率图(图 3(a)). 由图 3(a)易知, 北京市中心城区的 SARS 发病率明显要高于郊区, 发病率在东西方向的延展强度要高于南北方向, 发病率的整体分布与北京市的交通环线有一定的对应关系.

2.2 发病率图的 Bayesian 调整

根据空间统计单元上的发病者数量除以对应人口数得到的发病率图并不能有效反映发病风险, 显然, 区域内的人口数越多, 计算得到的发病率越可靠, 反之, 区域内的人口数越少, 发病率值受随机性的影响越大, 也就越不可靠. 实际中, 人口数越多的地方, 行政分区划分越细密, 人口数越少的地方, 行政分区划分越稀疏, 于是, 直接计算得到的发病率图容易造成视觉假象. 因为人口数稀少但空间范围很广的区域显示的发病率结果存在更大的不确定性, 而那些更加重要且结果更可靠的发病率显示则由于行政分区过小而无法传递对等的正确信息, 因此, 直接计算的发病率图显示使不确定性越大的分区表现更为显著.

北京市 SARS 流行的平均发病率约为 21.7/100000. 对比分析北京市的 SARS 发病率与人口密度(图 4(a)), 可以发现人口密度小的区域的 SARS 发病率更加离散, 发病率值最高与最低的区域均处在人口密度小的区域, 且存在许多发病率为零的情况, 这显然不能真实反映 SARS 流行传播的潜在风险, 因此需要对直接计算的 SARS 发病率进行 Bayesian 调整, 使调整后的发病率图更能反映所在区域的 SARS 传播风险(图 3(b)). 相对于直接计算的发病率图(图 3(a)), Bayesian 调整后的结果使无 SARS 发病和发病率非常低的周边郊区地带(人口密度较低)的发病率有了一定的提高, 并降低了由于随机性而导致的高发病率值. Bayesian 方法对人口密度很高的城区地带的发病率的调整非常小, 因为这些区域的发病率值

是比较可靠的, 主要调整的是那些人口密度小且发病率显著偏离北京市平均发病率的城郊与远郊地带. 下面简要介绍 Bayesian 调整的实现方法^[26].

设分区 i 内真实的 SARS 发病率值 θ_i 是服从正态概率分布 $N(\mu_i, \sigma_i^2)$ 的随机变量, 计算得到的 SARS 发病率 t_i 只是正态概率分布 $N(\mu_i, \sigma_i^2)$ 的一次实现, Bayesian 调整结果是 SARS 发病率的观察值 t_i 与期望值 μ_i 的组性组合:

$$\hat{\theta}_i = w_i \cdot t_i + (1 - w_i) \cdot \mu_i, \quad (1)$$

$$w_i = \frac{\sigma_i^2}{\sigma_i^2 + \mu_i / n_i}, \quad (2)$$

其中, w_i 是 $[0, 1]$ 之间变化的权值, n_i 为分区内的人口数, 由式(1)和(2)知, n_i 越小, w_i 也越小, 从而会使 μ_i 所占的权重越大, $\hat{\theta}_i$ 更接近于期望值 μ_i , 反之, $\hat{\theta}_i$ 更接近于计算值 t_i .

只要估算得到了每个分区正态概率分布的期望值 μ_i 和标准差 σ_i , 直接代入式(1)和(2)便可得到 SARS 发病率的 Bayesian 调整结果. μ_i 和 σ_i 的估计比较复杂, 最简单的一种处理方式是认为整个研究区域内所有分子的 μ_i 和 σ_i 均保持一致, 于是可以利用所有分区内的发病人数与人口数值估算得到各个分区的发病率的期望值 μ_i 和标准差 σ_i :

$$\mu_i = \hat{\mu} = \sum y_i / \sum n_i, \quad (3)$$

$$\sigma_i^2 = \frac{\sum n_i (t_i - \hat{\mu})^2}{\sum n_i} - \frac{\hat{\mu}}{\bar{n}}, \quad (4)$$

其中, $\hat{\mu}$ 为研究区域的平均 SARS 发病率, \bar{n} 为研究区域各分区的平均人口数.

2.3 SARS 发病率图的空间平滑

受随机性的影响, 根据直接计算得到的发病率图有时很难观察到明显的空间模式(见图 3(a)), 高低发病率值在空间上交错分布, 如果不仔细观察, 很难观察到发病率在空间上的分布特征与趋势. 空间平滑方法可以有效解决这一问题, 主要是利用遍历全部空间统计单元的移动框内所有发病率的某一特征值代替直接计算的发病率值, 特征值可以有算术均值、几何均值、中间值等, 空间平滑方法的思想是牺牲部分细节信息, 突显宏观的空间趋势. 本文采用算术均值来对北京市 SARS 发病率进行空间平滑.

影响空间平滑结果最重要的因素是移动框的定

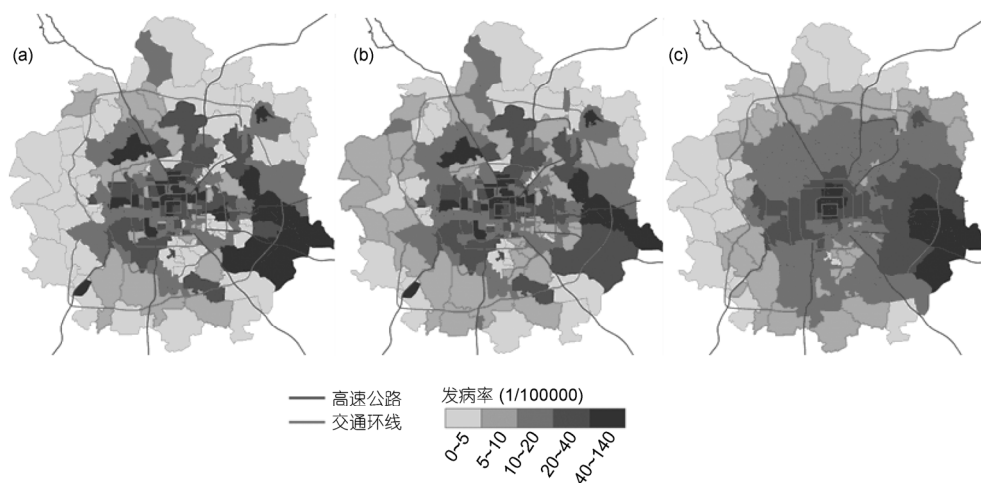


图3 SARS 发病率图

(a) 直接计算的发病率图; (b) Bayesian 调整后的发病率图; (c) 空间平滑后的发病率图

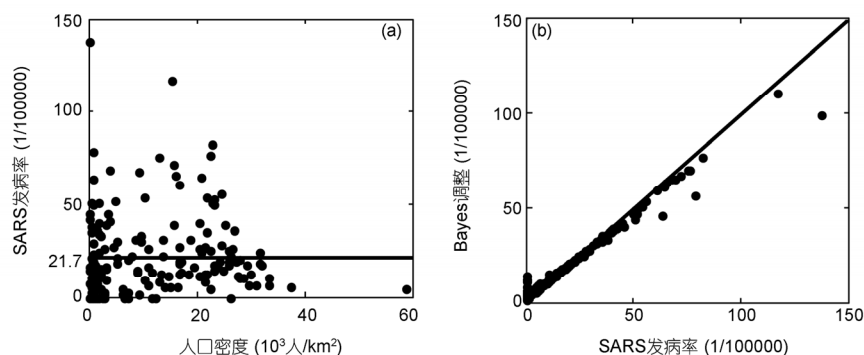


图4 SARS 发病率的 Bayesian 调整

(a) SARS 发病率与人口密度的对比; (b) Bayesian 调整前后的对比

义方式和大小. 移动框可以是矩形方框、圆框或自适应的空间邻近范围, 论文中选用空间邻接矩阵来定义移动框, 这主要考虑到传染病的扩散传播特征会使子区内的发病率与邻近子区的发病率趋向类同.

空间邻接矩阵是一个二元对称的空间权重矩阵, 它主要用来定量表达空间对象的邻近关系, 其表达方式为

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}, \quad (5)$$

其中, w_{ij} ($i=1, 2, 3, \dots, n; j=1, 2, 3, \dots, n$) 为表示对象 i ,

j 空间邻接关系的特征值.

定义空间对象邻接矩阵的方法有很多, 主要分为两大类: 拓扑与距离. 不同空间邻接矩阵的定义方式发病率的空间平滑结果有一定影响, 本文中选用一阶拓扑的空间邻接矩阵来定义空间平滑的移动框. 对 2003 年北京市的 SARS 发病率进行空间平滑(图 3(c)), 可以清楚地看到北京市中心城区的 SARS 发病率要明显高于郊区地带, 从城市中心到城郊的过渡中, 发病率值逐渐减小, 呈现出与道路交通环线类似的环状渐变特征. SARS 发病率在北京市东部形成了一个异于一般渐变模式的局部高风险地带, 表明 2003 年北京 SARS 流行期间在北京市东部存在一个

异于其他区域的 SARS 传播通道.

3 SARS 流行的时空传播规律与影响因素分析

3.1 空间相关性的定量计算方法

地理空间对象普遍存在自相似性, 距离越近相似程度越高^[27], 它是自然界存在秩序、格局和多样性的根本原因之一^[28]. 传统的相关性分析主要用于定量描述两组事物的共变趋势, 其理论基础是样本独立的经典统计学方法. 然而, 传染病在空间上的传播扩散与人口分布、环境及其他各类空间因子息息相关, 感染者之间非完全独立, 有一定的空间依赖性, 因而经典的相关性分析方法并不适用于传染病空间特征的分析. 现代空间统计分析方法正是基于样本非独立的地理空间对象发展起来的强大分析工具, 与普通相关性分析用于定量描述两组对象共变性特征不同, 空间相关性分析主要定量表达事物本身的空间关联关系, 即对象与邻近对象之间的共变特征.

用于定量衡量地理空间对象空间自相关特征的数学表达方法有很多, 基于全局空间自相关特征表达的指标主要有: Moran's I, Geary's C, Getis's G 和 Join Count 等^[29], 基于局部空间自相关特征表达的指标主要有: Local Moran's I, LISA(Local Indicators of Spatial Association)^[30]和 Getis's G^{*}^[31], 另外还有地统计学(Geostatistic)中常用的半变异图(Semivariogram)^[32]. 本文主要利用半变异函数(Semivariogram)、Moran's I 统计指数和 LISA 统计指数对北京 SARS 发病率的空间传播风险进行定量分析.

3.1.1 Semi-variogram

地统计学中的半变异函数主要根据有限的样本点数值来定量描述空间对象的变异大小(或空间相似性大小), 其计算过程主要分为两步, 第一步是根据已有的样本点数值计算经验半变异值, 第二步是采用数学模型对经验半变异值进行拟合, 得到理论半变异函数.

(1) Semi-variogram 的经验计算公式^[32]:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \sum_{|s(x_i)-s(x_j)| \in h \pm \varepsilon} [z(x_i) - z(x_j)]^2, \quad (6)$$

其中, $z(x_i)$, $s(x_i)$ 分别为第 i 个样本点的样本值和空间

位置, w_{ij} 为样本点 i, j 的空间邻接特征, $N(h)$ 为距离 $h \pm \varepsilon$ 内的样本点对的数量. Semi-variogram 可用于传染病传播的驱动力分析.

(2) Semi-variogram 3 个最常用的理论计算公式^[32]:

1) 球状模型:

$$\gamma(h) = \begin{cases} 0, & h = 0, \\ C_0 + (C - C_0) \left(\frac{3h}{2a} - \frac{h^3}{2a^3} \right), & 0 < h \leq a, \\ C, & h > a. \end{cases} \quad (7)$$

2) 高斯模型:

$$\gamma(h) = \begin{cases} 0, & h = 0, \\ C_0 + (C - C_0)(1 - e^{-3h^2/a^2}), & h > 0. \end{cases} \quad (8)$$

3) 指数模型:

$$\gamma(h) = \begin{cases} 0, & h = 0, \\ C_0 + (C - C_0)(1 - e^{-h/a}), & h > 0. \end{cases} \quad (9)$$

其中, $\gamma(h)$ 为相距 h 的两点之间的变异值, C_0 (Nugget) 为块金值, 表示随机性大小, 它主要是由于研究区域本身固有的随机性及样本数据无法反映小于最小样本点距离的变异性所致, C (Sill) 为基台值, 它与研究区域内独立样本点的方差一致, a (Range) 为影响范围, 表示两点之间具有空间相关性的阈值, 当两点之间的距离大于此阈值时, 两点之间将相互独立.

3.1.2 Moran's I 统计指数

Moran's I 主要用于整体上衡量空间对象之间的相关性, 其计算公式^[33]为

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (10)$$

其中, x_i, x_j 为空间上 i, j 两点的属性值, \bar{x} 为研究区域内所有样点属性值的平均值; w_{ij} 为空间统计单元 i 与 j 的空间邻接特征值, 用于表示空间步长, 一般采用拓扑关系的邻接矩阵, 即 i, j 在空间上存在边界线的共享, 则 $w_{ij}=1$, 否则为 0; n 为研究区域内所有空间统计单元的总数.

Moran's I 的取值范围为 $[-1, 1]$, 正值表示对象在空间上正相关, 负值则表示负相关, Moran's I 的期望值为 $-1/(n-1)$, 表示完全随机, Moran's I 的统计学意义

可通过标准正态化后的 Z-Score 检验来评价。

3.1.3 LISA 统计指数

Moran' I 指数只能从整体上对整个区域的空间相关性进行平均度量, 它无法反映空间异质情况时各个分区空间聚类情况。1995 年美国区域经济学家 Anselin^[30]提出了用于度量局部空间聚类特征的 LISA (Local Indicators of Spatial Association)统计指标:

$$I_i = \frac{n(x_i - \bar{x}) \sum_{j=1}^n w_{ij} (x_j - \bar{x})}{\sum_{j=1}^n (x_j - \bar{x})^2}, \quad (11)$$

其中, x_i, x_j 与 w_{ij} 的含义同式(10)中的定义。

3.2 SARS 传播的空间变异特征

由于许多不确定性因素的综合影响, 北京市 SARS 发病率出现了较多破碎情况(图 3(a)), 这会影响到北京市 SARS 发病率的空间变异特征的有效识别。因此, 论文中采用空间平滑后的北京市 SARS 发病率作为空间变异特征分析的对象(图 3(c))。

采用半变异函数来描述空间变异特征需要地理空间对象服从二阶平稳性的条件, 即研究区域内任何一点的随机变量的期望值相等, 且地理空间中任意两点的变异大小(或协方差)仅与两点的空间距离与方位有关, 而与这两点所处的空间位置无关。显然, 北京 SARS 发病率由城区中心向外发病率逐渐降低的空间模式与二阶平稳中的期望值相等的条件不符(见图 3(c)), 因此需要预先进行去趋势面的转换。二阶趋势面能够较好的反映由城市中心向城郊渐变

中 SARS 发病率逐渐衰退的整体趋势, 故论文中采用对 SARS 发病率的空间平滑值进行去二阶趋势面的预处理方法, 使去全局二阶趋势面后北京市各分区的发病率的期望值近似相等。

采用软件 ArcGIS9.2 的地统计工具来分析 2003 年北京市 SARS 空间传播风险的变异情况。由 SARS 发病率的半变异图(图 5)与半变异函数的参数估计结果(表 2), 易知北京市 SARS 的流行传播具有强的空间自相似性, 且在空间上的扩散存在显著的异质性, 东北-西南方向(偏向东西方向)的空间变异最小, 空间影响距离约为 13.627 km, 西北-东南方向(偏向南北方向)的空间变异最大, 空间影响距离约为 8.545 km。变异越大, 表明 SARS 在空间上的扩散速度越慢, 反之, 扩散速度越大, 因此, 根据图 5(b)的研究结果, 可以发现 2003 年北京市的 SARS 流行中, 城市中心向城郊的空间扩散在西北-东南(偏南北向)的速度明显弱于东北-西南(偏东西向)方向, 故东北-西南方向(偏东西向)的人群受 SARS 传播的威胁更大, 其原因有可能是西北-东南方向的人群活动更频繁所致。

空间变异特征的识别结果能够用于指导北京市对传染病的防控措施在空间上的优化布局, 对于 2003 年北京市 SARS 流行情况而言, 根据图 5(b)中的椭圆形状进行层级防控会是一种好的优化选择, 这

表 2 半变异函数的参数估计

影响半径(km)		方位角 (°)	块金值 C_0	基台值 C	C_0/C
长轴方向	短轴方向				
13.627	8.545	74.70	29.424	188.574	0.156

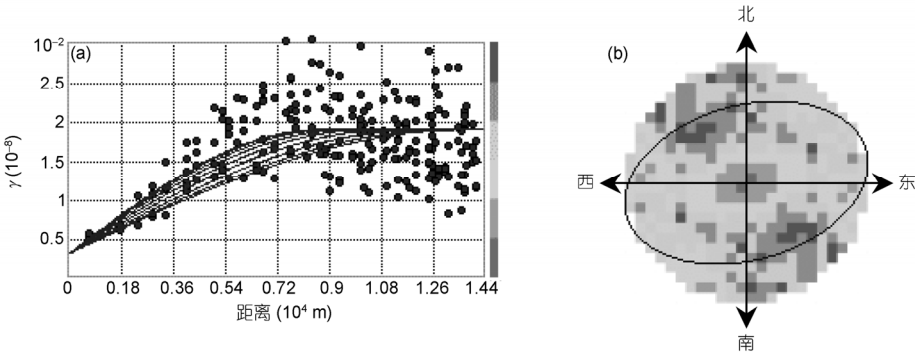


图 5 SARS 传播的半变异图

(a) 经验半变异图; (b) 半变异表面

对未来有可能重现的 SARS 或其他在人群间急性传播的传染病的科学防控有重要借鉴意义。

3.3 SARS 传播的全局空间相关性特征

利用空间分析软件 GeoDa (<https://www.geoda.uiuc.edu>)计算得到 2003 年北京市的 SARS 发病率 (Bayesian 调整后)的 Moran's I 统计指数为 0.2985。零假设 H_0 : 发病率在空间上完全随机分布, 根据零假设 H_0 进行 999 次 Monte Carlo 随机模拟得到 Morna's I 的期望值 $E(I)=-0.0052$ 、标准差 $S_d=0.0443$, 由此计算得到 z-score 为 6.74, $P<0.01$, 表明整体上北京市 SARS 发病率在空间上显著正相关。

性别、年龄、发病时间及接触方式等都会对 SARS 的空间传播模式产生影响, 论文采用上面同样的计算方法对这些影响因素进行考察, 结果如表 3 所示。由表 3 知: ① 女性感染者的空间相关性强于男性感染者, 这可能因为男性的活动范围要大于女性, 致使男性感染在空间上的分布更发散; ② 有过接触史的感染者的空间相关性明显强于没有接触史的感染者, 这可能因为有接触史感染者的感染途径几乎全部源于医院传播和家庭传播, 而无接触史感染者的感染途径则比较随机和分散(如餐馆吃饭、乘电梯、

公园游玩、商场购物等公共场所), 因此有接触史的感染者在空间上更聚集; ③ 年龄越大, 空间相关性更强, 这是由于年龄越大的人受医院或家庭传播的影响越大所致。如图 6 所示, 随着年龄值的增加, 有/无接触史感染人数的比值逐渐增大, 表明年龄越大的人受医院/家庭传播途径的影响越强; ④ SARS 感染率的空间相关性经历了一个由弱到强再到弱的变化过程, 这可以解释为: 早期 SARS 感染人数较少, 新发病例受感染的途径较多, 故 SARS 传播较为发散, 自相关性不强, 在 SARS 流行高峰期, 人群活动受到政府应急措施的限制, 故这一时期的新发感染大多与医院和家庭传播有关, 空间聚集性较强, SARS 流

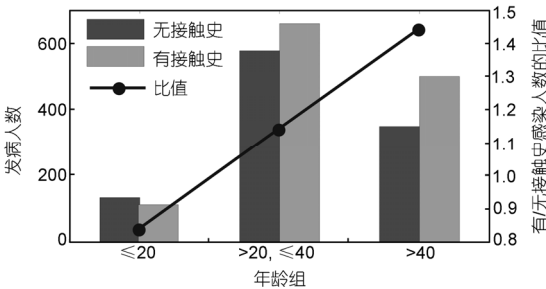


图 6 年龄与接触史的关系

表 3 不同类型人群的 Moran's I 统计结果^{a)}

SARS 感染者		Moran's I	999 次完全随机分布的模拟结果		
			期望值	标准差	Z-score
性别	男	0.2231*	-0.0020	0.0448	4.9799
	女	0.2906*	-0.0025	0.0418	6.9522
是否有过接触史	有	0.3406*	-0.0019	0.0405	8.4099
	非	0.1317*	-0.0039	0.0517	2.5474
年龄	≤20	0.1291*	-0.0031	0.0446	2.8946
	>20, ≤40	0.1708*	-0.0013	0.0469	3.6418
	>40	0.3667*	-0.0013	0.0394	9.3071
时间 2003-03-08~ 2003-05-31	第 1 周	0.0158	-0.0037	0.0237	0.6667
	第 2 周	0.0731*	-0.0012	0.0252	2.9008
	第 3 周	-0.0104	-0.0058	0.0383	-0.2715
	第 4 周	0.0996*	-0.0055	0.0385	2.5870
	第 5 周	0.1713*	-0.0088	0.0471	3.6369
	第 6 周	0.2217*	-0.0036	0.0405	5.4741
	第 7 周	0.1953*	-0.0041	0.0475	4.1116
	第 8 周	0.2368*	-0.0013	0.0418	5.6651
	第 9 周	0.0739	-0.0048	0.0431	1.7146
	第 10 周	0.1162*	-0.003	0.0426	2.7277
	第 11 周	0.1997*	-0.0049	0.0416	4.8005
	第 12 周	0.0463	-0.0039	0.0285	1.6246

a) *表示存在显著的空间正相关($P<0.05$)

行晚期, 医院和家庭传播途径已得到有效控制, 此时零星的新发病例大多感染途径不明, SARS 出现在空间上又变得较分散。

3.4 SARS 传播的局部相关性特征与时空传播规律

采用 LISA 指数来分析 2003 年北京市 SARS 流行的空间传播风险及其时空演化特征, LISA 指数能够定量反映出空间上的 4 种聚类特征: ① 高值聚集(高-高)-高风险值的空间统计单元邻近的周围区域仍然是高值; ② 低值聚集(低-低)-低风险值的空间统计单元邻近的周围区域仍然是低值; ③ 高低聚集(高-低)-高风险值的空间统计单元邻近的周围区域是低值; ④ 低高聚集(低-高)-低风险值的空间统计单元邻近的周围区域是高值。高值聚集为热点区域, 表示传

染病以高于平均流行的水平在空间上聚集传播, 因此处于这些区域的人群受 SARS 感染的风险非常高。与此相反, 低值聚集为冷点区域, 表示这些区域的传染病处于低风险传播水平, 高低聚集表示从高风险传播区域向低风险传播区域的过渡, 低高聚集表示从低风险传播区域向高风险传播区域的过渡, 不属于这 4 种聚类的空间范围均为随机区域。

利用空间分析软件 GeoDA 对 Bayesian 调整后的 SARS 发病率进行基于 LISA 统计指标的传播风险分析(图 7(a)), 可以发现北京 SARS 发病的高风险(热点区域)主要集中在城市中心(三号环线以内的北部区域)与东部稍偏南的城乡交接地带(东六环与通州交界的区域), 北京市西北方向城郊区域的 SARS 传播风险较低, 相对较为安全, 北京市四环、五环线附近

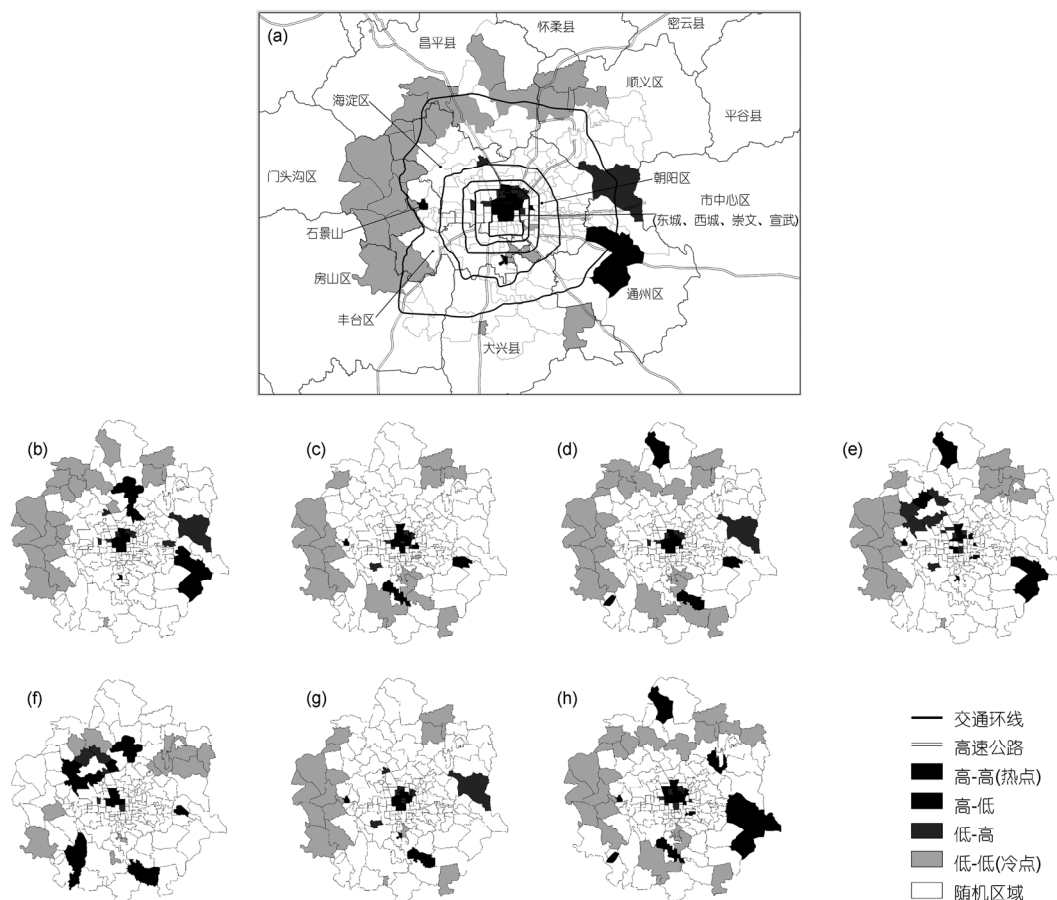


图 7 SARS 传染病的空间传播风险

(a) 所有人群; (b) 男性感染人群; (c) 女性感染人群; (d) 有过接触史的感染人群; (e) 无接触史的感染人群; (f) 年龄 ≤ 20 岁的感染人群; (g) 20岁 $<$ 年龄 ≤ 40 岁的感染人群; (h) 年龄 > 40 的感染人群

区域的 SARS 传播风险呈随机分布. 这一实验结果可以得到几点 SARS 防控的经验: ① 北京市中心区域发病风险显著偏高, 一方面是由于这个区域的人口密度高, 人们活动交流的频率大, 另外一个重要原因可能与 SARS 传播早期这个区域收治的 SARS 患者的医院众多有关, 医院传播途径导致了 SARS 收治医院附近区域的高风险传播; ② 北京市东部偏南区域的人口密度并不高, 且接收 SARS 感染者的医院也不多, 何以在这一区域出现局部的高传播风险? 其原因有可能与 CBD 商圈有关, 因为 CBD 商圈地带的生活成本很高, 许多在 CBD 工作的人选择在生活成本较低且靠近 CBD 的通州区租房居住, 形成了由城市中心到通州区的 SARS 传播通道; ③ SARS 流行期间的防控措施在西北方向的效果更显著, 从而使西北方向的正常人群受到了更好保护; ④ 整体来看, 北京市 SARS 防控的效果是非常显著的, 使 SARS 高风险传播的区域局限在非常有限的城市中心地带, 没有形成大范围扩散.

根据性别、是否有过接触史及年龄 3 个因素, 分别考察不同类型人群在空间上传播扩散的异质性特征(图 7(b)~(h)). 由图可知: 性别对 SARS 传播的空间风险的影响非常小; 有过接触史和无接触史的 SARS 感染人群的空间相关性特征存在一些差异, 无接触史感染人群的空间热点的出现除在城市中心和东南

城郊外, 还在北京市西北方向有局部出现; 年龄在 20~40 之间的中青年感染者的空间热点仅在北京市中心区有出现, 而年龄低于 20 岁和高于 40 岁的少、老年人的空间热点除在城市中心区外, 还分别在城市西北郊和东南郊有出现, 这表明防控措施对中青年的效果更显著, 相对而言, 老、少年龄段人群的防控效果较差.

根据 2003 年北京市 SARS 流行在不同时期的空间传播风险情况 (图 8), 可以探寻 SARS 流行的热点区域随时间的演化规律. 根据图 8 并结合各个时期的重要措施对 SARS 时空传播规律进行流行病学解释如表 4 所示.

由表 4, 可以得到如下几点重要启示: ① SARS 传染病的危害极大, 若无人干预, 传染病流行导致的被感染人数将呈指数级增长, 因此, 在疫情早期实施快速有效干预是降低 SARS 危害的关键, 3 月份是将北京 SARS 流行消灭在萌芽状态的关键时段, 这一时期对 SARS 危害认识的不足及应急反应的迟缓是直接导致 2003 年北京 SARS 大范围流行的主要因素; ② 北京市 SARS 高风险传播的热点区域面积经历了一个少数零星-多数集中-少数零星的变化过程, SARS 流行在北京有两个高风险聚集的热点区域——城市中心和东部郊区, 城市中心的热点区域经历了由零星发散到大范围聚集再到零星发散的过程, 东

表 4 2003 年 SARS 流行的时空传播规律及其流行病学解释

时间	SARS 流行的时空传播规律		应急政策与措施	流行病学解释
	时间特征	空间特征		
第 1~3 周 (3-8~3-28)	新发病例较少, 增长速度平缓	热点区域很少, 在城市中心地带呈零星发散分布	SARS 防控仅限于医院, 由于认识不足, 导致了大量医务人员感染	SARS 经由香港登陆北京, SARS 流行处于早期酝酿状态, 直接危害有限, 但潜在威胁巨大, 是科学防控的关键时期
第 4~5 周 (3-29~4-11)	新发病例逐渐增多, 增长速度加快	热点区域增长迅速, 呈聚集性状, 且开始由城市中心向外围扩散, 尤其是向东部	WHO 向中国发出 SARS 流行警告	SARS 开始转向大范围扩散传播, 危害加剧, 高风险区域向北京市东部扩散, 并在东部形成高风险的聚集区域, 北京市西北方向的扩散得到有效遏制
第 6~8 周 (4-12~5-2)	新发病例急剧增加, 并达到最高峰	正式形成城市中心和东部两个热点聚集的高风险区域, 中心热点的增长得到有效控制, 东部热点呈持续增长态势	全民动员抗击“非典”, SARS 信息畅通, 确定非典定点医院, 交叉感染严重的医院被整体隔离, 停止娱乐场所经营, 管制交通及人群集体活动	SARS 流行全面暴发, SARS 危害和疫情信息公开, 正确防护措施得以广泛宣传. SARS 危害虽大, 但已开始受到有效控制, 每周新发病例开始下降, 没有形成进一步的空间扩散
第 9~10 周 (5-3~5-31)	新发病例迅速下降, SARS 流行逐渐平息	东部热点区域迅速变小并消亡, 城市中心热点区域也在逐渐变小, 热点有向西南和西北方向的转移, 但迅速得到控制	建立小汤山定点医院, 集中收治 SARS 感染者, 宣传非典及其防治措施, 开展公共卫生建设, 疫情对民众全面公开, 继续管制公众场所、交通及人群集体活动	防控效果开始全面显现, 疫情得到有效控制, 小汤山定点医院的设立基本杜绝了 SARS 的医院传播途径, “早发现、早隔离、早防治”的全面贯彻使 SARS 的空间扩散得到全面遏止

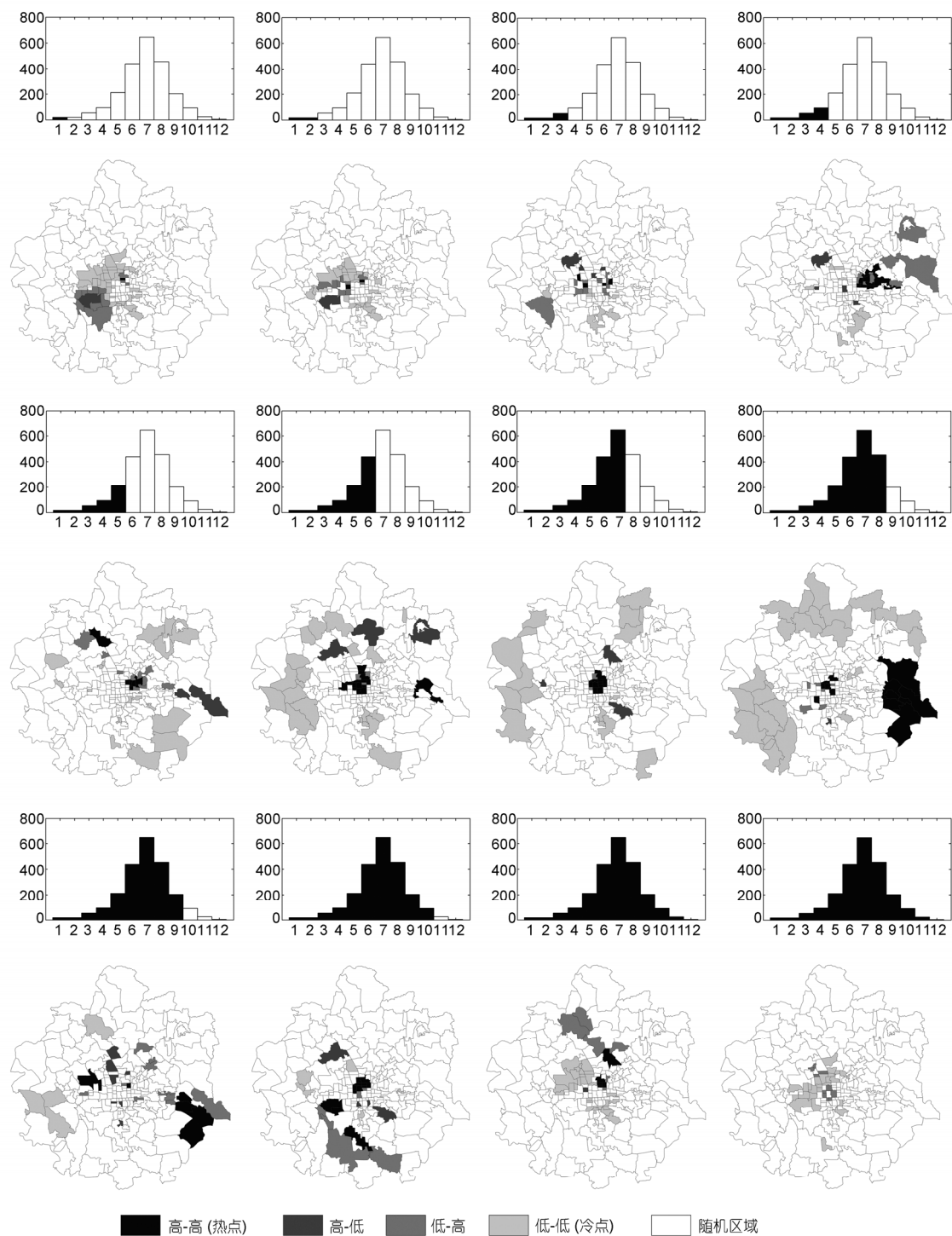


图 8 SARS 传播风险的时空演化过程

共有 3 组柱状图和地图。柱状图是每周新增 SARS 感染人数的时序变化，横轴是时间，单位：周，纵轴是每周的新发病例数量，单位：人；地图是对应时段内 SARS 传播的空间聚类特征

部热点区域在 4 月初形成, 中间虽有消长, 但在随后一个多月的时间内没有得到有效控制, 这种状况一直延续到 5 月中旬; ③ 高风险的热点区域始终局限于有限的局部范围, SARS 的空间扩散得到了较好遏止, 尤其是在西北方向; ④ 4 月 20 日后采取的一系列措施收到了非常好的效果, 尤其是小汤山医院的设立, 不仅使新增 SARS 发病人数急剧下降, 而且发病热点区域迅速缩小, 没有形成再次空间扩散。

4 总结与讨论

流行病学主要研究疾病在人群中的分布规律以及其影响因素, 用于指导科学有效的防控措施, 流行病学研究是传染病防控的科学基础, 对保障公共卫生安全有重要作用。流行病学是一项投入较少, 但成效极为显著的重要学科, 新发突发传染病流行时, 特异性疫苗与根本性治疗方法一般需要很长时间, 此时, 有效应急防控的措施通常是隔离、追踪密切接触者、管制口岸与交通、限制人群活动等, 如何根据有限的医疗及社会资源来实现科学防控需要流行病学研究成果的支持。相对于国外而言, 我国的流行病学研究长期滞后, 尤其是对于传染病时空传播规律的研究。时间、空间和属性是事物信息的 3 个基本要素, 传统的基于经典统计学的描述性流行病学方法虽然能够揭示许多传染病在人群中的流行规律, 但缺乏对空间信息及时空联合信息的有效反映, 无法反映

传染病的时空传播规律。

2003 年的 SARS 流行事件暴露了我国在流行病学研究方面的不足。时至今日, 虽已有大量针对 SARS 流行规律的研究, 但大多基于 SARS 疫情的趋势预测与人群影响, 对 SARS 流行的空间扩散过程、时空传播规律及其影响因素的认识仍然不清楚。本文以 2003 年北京市 SARS 感染者的时空数据为基础, 利用地理信息技术与空间统计方法对北京 SARS 流行的空间模式、时空传播规律及其影响因素进行了全面研究, 并对政策措施的得失进行了评价。研究成果不仅有助于更深刻地认识 SARS 流行规律, 也可为未来可能重现的 SARS 流行及其他新发突发传染病的科学防控提供有益帮助。

近 20 年来, 随着信息技术的发展与广泛应用, 海量时空数据的存储、管理与动态更新成为可能, 基于时空数据的现代空间统计分析方法得到迅猛发展, 并已在资源、环境、生态、经济、社会等领域得到广泛应用。在公共健康方面, 现代空间统计分析已取得令人瞩目的成功, 目前, 国外已将公共健康与空间统计分析方法相结合形成了一个专门学科——空间流行病学, 它能更深入地发掘传染病流行的时空规律及其驱动力, 这对全面认识传染病流行的内在规律与影响因素有重要意义。这方面的研究在西方发达国家普遍受到重视并广泛应用于各类传染病流行事件。目前, 我国在这方面仍处于起步阶段, 本文研究方法能够为流行病规律的研究提供很好的借鉴。

致谢 感谢审稿专家提出的宝贵意见。

参考文献

- 1 Oxford J S, Bossuyt S, Lambkin R. A new infectious disease challenge: Urbani severe acute respiratory syndrome (SARS) associated coronavirus. *Immunology*, 2003, 109: 326—328
- 2 Marra M A, Jones S J, Astell C R, et al. The genome sequence of the SARS-associated coronavirus. *Science*, 2003, 300: 1399—1404
- 3 Tan Y M, Chow P K H, Soo K C. Severe acute respiratory syndrome—Clinical outcome after inpatient outbreak of SARS in Singapore. *Br Med J*, 2003, 326: 1394
- 4 Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 2003, 300: 1966—1970
- 5 Keogh-Brown M R, Smith R D. The economic impact of SARS: How does the reality match the predictions? *Health Policy*, 2008, 88: 110—120
- 6 Chen W K, Wu H W, Lin C C, et al. Emergency department response to SARS, Taiwan. *Emerging Infect Dis*, 2005, 11: 1067—1073
- 7 Pang X H, Zhu Z H, Xu F J, et al. Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing, 2003. *JAMA*, 2003, 290: 3215—3221
- 8 Liang W N, Liu M, Chen Q, et al. Assessment of impacts of public health interventions on the SARS epidemic in Beijing in terms of the

- intervals between its symptom onset, hospital admission, and notification. *Biomed Environ Sci*, 2005, 18: 153—158
- 9 梁万年, 刘民, 陈琦, 等. SARS 患者发病及病例报告时间干预效果分析. *中国公共卫生*, 2004, 20: 129—130
- 10 Zhang Z B. The outbreak pattern of SARS cases in China as revealed by a mathematical model. *Ecol Model*, 2007, 204: 420—426
- 11 汪海英, 荣峰, 柯孚久, 等. 严重急性呼吸综合征(SARS)的传播与控制动力学. *科学通报*, 2003, 48: 1933—1935
- 12 蔡全才, 姜庆五, 徐勤丰, 等. 定量评价 SARS 干预措施效果的传播动力学模型. *中华流行病学杂志*, 2005, 26: 153—158
- 13 韩卫国, 王劲峰, 刘旭华. SARS 传播时间过程的参数反演和趋势预测. *地球科学进展*, 2004, 19: 925—930
- 14 宗跃光, 王莉, 曲秀丽. 基于蒙特卡罗模拟法的北京地区非典时空变化特征. *地理研究*, 2004, 23: 815—824
- 15 石耀霖. SARS 传染扩散的动力学随机模型. *科学通报*, 2003, 13: 1373—1377
- 16 陈奇志. 随机模型在非典型肺炎预测及疫情分析中的应用. *北京大学学报(医学版)*, 2003, 35: 75—80
- 17 Wu J, Xu F J, Zhou W G, et al. Risk factors for SARS among persons without known contact with SARS patients, Beijing, China. *Emerging Infect Dis*, 2004, 10: 210—216
- 18 王铮, 蔡砥, 李山. 中国 SARS 流行的季节性风险探讨. *地理研究*, 2003, 22: 541—550
- 19 Li R W K, Leung K W C, Sun F C S, et al. Severe Acute Respiratory Syndrome (SARS) and the GDP. Part I: Epidemiology, virology, pathology and general health issues. *Br Dent J*, 2004, 197: 77—80
- 20 朱迎波, 葛全胜, 魏小安. SARS 对中国入境旅游人数影响的研究. *地理研究*, 2003, 22: 551—559
- 21 Meng B, Wang J F, Liu J Y, et al. Understanding the spatial diffusion process of severe acute respiratory syndrome in Beijing. *Public Health*, 2005, 119: 1080—1087
- 22 Wang J F, McMichael A J, Meng B, et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. *Bull World Health Organ*, 2006, 84: 965—968
- 23 王劲峰, 孟斌, 郑晓瑛, 等. 北京市 2003 年 SARS 疫情的多维分布及其影响因素分析. *中华流行病学杂志*, 2005, 26: 164—168
- 24 武继磊, 王劲峰, 孟斌, 等. 2003 年北京市 SARS 疫情空间相关性分析. *浙江大学学报(农业与生命科学版)*, 2005, 31: 97—101
- 25 Liu M, Liang W N, Du H, et al. Epidemiological features of severe acute respiratory syndrome in Beijing urban and suburb areas in 2003. *Biomed Environ Sci*, 2005, 18: 227—232
- 26 Anselin L, Kim Y W, Syabri I. Web-based analytical tools for the exploration of spatial data. *J Geogr Syst*, 2004, 6: 197—218
- 27 Tobler W. A computer movie simulating urban growth in the detroit region. *Econ Geogr*, 1970, 46: 234—240
- 28 Goodchild M F. The application of advanced tecnology in assessing environmental impacts, In: Corwin D L, Loague K, eds. *Application of GIS to the Modeling of Non-point Source Pollutants in the Vadose Zone*. *Soil Sci Soc Am Spec Publ*, 1996, 48: 1—17
- 29 Haining R P. *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University, 2003
- 30 Anselin L. Local indicators of spatial association-lisa. *Geogr Anal*, 1995, 27: 93—115
- 31 Ord J K, Getis G. Local spatial autocorrelation statistics: Distributional issues and application. *Geogr Anal*, 1995, 27: 286—306
- 32 Cressie N. *Statistics for Spatial Data*. New York: Wiley, 1993
- 33 Moran P. The interpretation of statistical maps. *J R Stat Soc Ser B-Stat Methodol*, 1948, 10: 243—251