

# 前瞻性时空重排扫描统计量法在深圳市流行性腮腺炎聚集性疫情早期预警中的应用

周志峰<sup>1</sup>, 廖玉学<sup>2</sup>, 李学云<sup>1</sup>, 尹凌<sup>3</sup>, 许玉成<sup>1</sup>, 梁静<sup>1</sup>, 梅树江<sup>2</sup>

1. 深圳市福田区疾病预防控制中心, 广东 深圳 518040 2. 深圳市疾病预防控制中心, 广东 深圳 518055  
3. 中国科学院深圳先进技术研究院, 广东 深圳 518055

**摘要:** **目的** 分析前瞻性时空重排扫描统计量法应用于深圳市流行性腮腺炎疫情早期预警的效果, 为传染病监测预警系统的完善提供科学依据。 **方法** 以深圳市为研究范围, 以街道行政区域为研究尺度, 使用 SaTScan 软件前瞻性时空重排扫描模型对深圳市 2012 年 1 月 1 日—12 月 31 日流行性腮腺炎发病情况进行逐日模拟预警。根据国家疾病监测信息报告管理系统 (大疫情系统) 上报的病例个案信息、深圳市疾病控制信息管理系统 (深圳系统) 上报的聚集性疫情事件, 核实模拟预警信息的真实性, 计算模型的灵敏度、错误预警率、暴发探测时间、预警提前时间等指标。比较时空重排扫描模型和国家传染病自动预警系统 (China Infectious Disease Automated-alert and Response System, CIDARS) 时间序列模型、时空序列模型预警效果的差异。 **结果** 深圳系统 2012 年共上报 29 宗流行性腮腺炎聚集性事件, 共计纳入深圳市 2011 年 12 月 3 日—2012 年 12 月 31 日的流行性腮腺炎分析病例 9 756 例, 累计分析 366 次, 其中  $P \leq 0.05$  的预警信号数为 88 条, 被确认为阳性信号的有 16 次, 模型的灵敏度为 55.17% (16/29), 错误预警率为 81.82% (72/88)。暴发探测时间在 0~22 d 之间, 平均 6.25 d, 中位数为 4 d。预警提前时间在 -12~41 d 之间, 平均 9.75 d, 中位数为 5.5 d。CIDARS 中 2012 年深圳市流行性腮腺炎时间序列预警信号发出 319 次, 被判断阳性的 2 次, 灵敏度为 6.90% (2/29), 错误预警率为 99.37% (315/317), 时空序列预警信号发出 208 次, 被判断阳性 3 次, 灵敏度为 10.34% (3/29), 错误预警率为 98.56% (205/208)。时空重排扫描模型信号数比大疫情系统时间序列模型和时空序列模型预警分别下降 72.41% 和 57.69%。时空重排扫描模型灵敏度显著高于大疫情系统中时间序列模型 ( $P < 0.001$ ) 和时空序列模型 ( $P < 0.001$ )。时空重排扫描模型的错误预警率显著低于 CIDARS 时间序列 ( $P < 0.001$ ) 和时空序列模型的错误预警率 ( $P < 0.001$ )。 **结论** 前瞻性时空重排扫描模型在深圳市流行性腮腺炎疫情中具有较好的早期预警效果。

**关键词:** 前瞻性时空重排扫描统计量; 流行性腮腺炎; 早期预警

中图分类号: R512.1 文献标识码: A 文章编号: 1006-3110(2020)01-0016-05 DOI: 10.3969/j.issn.1006-3110.2020.01.005

## Application of prospective spatiotemporal rearrangement scan statistic to early warning of mumps clustering epidemic in Shenzhen city

ZHOU Zhi-feng<sup>1</sup>, LIAO Yu-xue<sup>2</sup>, LI Xue-yun<sup>1</sup>, YIN Ling<sup>3</sup>, XU Yu-cheng<sup>1</sup>, LIANG Jing<sup>1</sup>, MEI Shu-jiang<sup>2</sup>

1. Futian District Center for Disease Control and Prevention, Shenzhen, Guangdong 518040, China;  
2. Shenzhen Municipal Center for Disease Control and Prevention, Shenzhen, Guangdong 518055, China;  
3. Shenzhen Institutes of Advanced Technology, Shenzhen, Guangdong 518055, China

Corresponding author: MEI Shu-jiang, E-mail: sjmei66@163.com

**Abstract:** **Objective** To analyze the effect of prospective spatiotemporal rearrangement scan statistic on the early warning of mumps epidemic in Shenzhen city, and to provide a scientific basis for improving infectious disease surveillance and warning system. **Methods** Taking Shenzhen city and the street administrative areas as the research scale, the prospective spatiotemporal rearrangement scan model of the SaTScan software was used to conduct daily simulated early warning on the incidence of mumps in Shenzhen city from January 1 to December 31 in 2012. According to the case information reported by the National Disease Surveillance Information Report Management System (hereinafter referred to as the pandemic system) and clustering epidemic events

**基金项目:** 国家自然科学基金项目 (编号: 41771441); 广东省医学科研基金项目 (编号: A2017233); 深圳市科技计划项目 (编号: JCYJ20170307105031703); 深圳市科技计划项目 (编号: JCYJ20170307164104491)

**作者简介:** 周志峰 (1981-), 男, 陕西宁强人, 硕士, 副主任医师, 主要从事公共卫生应急和空间流行病学研究工作。

**通信作者:** 梅树江, E-mail: sjmei66@163.com。

reported by Shenzhen Disease Control Information Management System (hereinafter referred to as the Shenzhen system), the authenticity of simulated early warning information was verified, and the sensitivity, error warning rate, outbreak detection time and advance warning time of the model were calculated. We compared the differences in warning effects of the spatiotemporal rearrangement scan model as well as the time series model and the space-time sequence model of China Infectious Disease Automated-alert and Response System (CIDARS). **Results** A total of 29 mumps aggregation events were reported by the Shenzhen system in 2012. A total of 9,756 mumps analysis cases and 366 times of cumulative analysis from December 3, 2011 to December 31, 2012 in Shenzhen city were enrolled in this study, of which the number of warning signals with  $P < 0.05$  was 88, and 16 times were identified as positive. The sensitivity of the model was 55.17% (16/29), and the error warning rate was 81.82% (72/88). The outbreak detection time ranged from 0 to 22 days, with an average of 6.25 days and a median of 4 days. The advance warning time ranged from -12 to 41 days, with an average of 9.75 days and a median of 5.5 days. The time series warning signals were sent out for 319 times, and the positive ones were judged for 2 times, with a sensitivity of 6.90% (2/29) and an error warning rate of 99.37% (315/317). The spatial and temporal sequence warning signals were sent out for 208 times and judged to be positive for 3 times, with a sensitivity of 10.34% (3/29) and an error warning rate of 98.56% (205/208). The number of warning signals in the spatiotemporal rearrangement scan model was 72.41% and 57.69% lower than that in the time series model and the space-time sequence model of the pandemic system, respectively. The sensitivity of the spatiotemporal rearrangement scan model was significantly higher than those of the time series model ( $P < 0.001$ ) and the space-time sequence model ( $P < 0.001$ ) of the pandemic system. The error warning rate of the spatiotemporal rearrangement scan model was significantly lower than those of the time series model ( $P < 0.001$ ) and the space-time sequence model of CIDARS ( $P < 0.001$ ). **Conclusions** The prospective spatiotemporal rearrangement scan model has a better early warning effect in mumps epidemic in Shenzhen city.

**Key words:** prospective spatiotemporal rearrangement scan statistic; mumps; early warning

2008 年我国启动了基于法定报告传染病的国家传染病自动预警系统(China Infectious Disease Automated-alert and Response System, CIDARS), 系统中部分重点传染病实施单病例预警, 其他病种采用时间模型和基于时间模型的时空模型预警<sup>[1]</sup>。系统运行十余年, 各级、各地工作人员对 CIDARS 开展了一些预警效果评价工作<sup>[2-7]</sup>, 研究普遍发现, 单病例预警效果稳定可靠, 而时间模型和时空模型存在信号数量庞大、信号阳性率低等缺点, 难以高效实现传染病早期自动预警。深圳市作为试点, 应用了时间模型和时空探测模型, 系统运行中同样存在信号阳性率极低的情况, 如 2018 年深圳市共收到时间序列预警信号 1 943 条, 被判为疑似事件有 7 条, 时间序列信号阳性率为 0.36%; 收到时空序列信号 648 条, 被判为疑似事件的有 1 条, 时空序列信号阳性率为 0.15%; 大量的无效信号给基层人员疫情核实工作带来很大困难。殷菲等<sup>[8]</sup>开展的相关研究表明前瞻性时空重排扫描统计量法能够有效的对传染病暴发做出早期预警, 其研究选用的麻疹病种发病数相对较少, 在其他发病率较高病种效果如何还有待证实, 不同城市行政区域的地理形状和面积存在较大差异, 该方法是否适用于深圳市还有待研究证实。为此, 本研究以流行性腮腺炎为例, 以深圳市为扫描范围, 以街道为扫描尺度, 利用前瞻性时空重排扫描模型对 2012 年深圳市流行性腮腺炎实现逐日模拟预警, 探讨该模型在深圳市流行性腮腺炎聚集性疫情早期预警

的效果, 并比较其与 CIDARS 中预警效果的差异, 为深圳市传染病预警系统进一步完善提供科学参考。

## 1 资料与方法

1.1 疾病信息数据 通过“疾病监测信息报告管理系统”(大疫情系统)收集整理在 2011 年 12 月 3 日—2012 年 12 月 31 日现住址在深圳市的流行性腮腺炎病例数据, 并剔除住址不能明确到街道的病例。通过 CIDARS 收集 2012 年 1 月 1 日—12 月 31 日流行性腮腺炎病种的时间序列和时空序列预警信号, 以及信号核实情况数据; 通过“深圳市疾病控制信息管理系统”(深圳系统)收集同期流行性腮腺炎聚集性疫情信息。

1.2 地理信息数据 通过深圳市国土局获得深圳市街道行政边界矢量地图, 考虑数据的前后一致性, 将福田区新成立的华强北街道和福保街道的病例划归原所属街道, 新成立的光明新区、龙华新区、坪山新区的病例修改街道编码划归原所属各街道, 本研究按 55 个街道进行统计分析。经纬度均以十进制数表示, 精确到小数点后 5 位。

1.3 前瞻性时空重排扫描统计量法的模拟预警 因为本次研究基于街道行政区划扫描, 但深圳市街道层面的实际人口数难以准确获取, 故本研究选用不使用实际人口数的前瞻性时空重排扫描统计量方法<sup>[8]</sup>。前瞻性时空重排扫描分析是在地图基础上建立一个时空二维圆柱体活动窗口, 每个圆柱表示一个可能的时

1.4 评价指标及方法 根据大疫情系统上报的病例个案信息、CIDARS 发出的时间序列和时空序列预警信号、深圳系统上报的聚集性事件等信息,核实模拟预警信息的真实性,计算灵敏度、信号阳性率、错误预警率、暴发探测时间和预警提前时间等指标<sup>[11-14]</sup>,比较时空重排扫描模型和 CIDARS 中时间和时空模型预警效果,其中三者间灵敏度、信号阳性率、错误预警率比较采用 $\chi^2$  检验,两两比较采用 Bonferroni 法进行调整, $P<0.05$  差异有统计学意义。灵敏度即实际为聚集性疫情而被模型正确地判定的百分比,反映模型识别聚集性疫情的能力。信号阳性率指阳性信号占有所有预警信号的百分比。错误预警率指非阳性信号占有所有预警信号的百分比。暴发探测时间指聚集性疫情开始时间至模型预警信号发出的时间间隔。预警提前时间为预警信号发出至聚集性疫情报告时间间隔。

2.1 CIDARS 预警信号分析 2012 年深圳市流行性腮腺炎共计发出预警信号 527 次(时间序列模型 319 次,时空序列模型 208 次),被判断阳性事件 5 次(时

2.2 时空重排扫描模型分析 在对 2012 年 1 月 1 日—12 月 31 日深圳市流行性腮腺炎开展逐日模拟预警时,共计纳入深圳市 2011 年 12 月 3 日—2012 年 12 月 31 日的流行性腮腺炎分析病例 9 756 例,共计分析 366 次,其中  $P \leq 0.05$  的预警信号数为 88 条。根据 2012 年大疫情系统上报的病例个案信息,深圳系统上报的 29 宗流行性腮腺炎聚集性事件信息,核实模拟扫描运算产生的预警信息的真实性,时空重排扫描模型发出的 88 次信号中,被判定为阳性事件的预警信号有 16 次,其中单一街道内的有 4 次,跨街道的有 12 次,跨区的有 1 次,预警信号空间范围包含的街道数在 1~7 个之间,平均 2.81 个,中位数为 2 个。被确认的 16 次预警信号空间范围累计包含街道和对应的流行性腮腺炎聚集性疫情发生地点见图 1。

图1 被确认的16条预警信号空间范围和对应的流行性腮腺炎聚集性疫情发生地点



信号阳性率 18.18% (16/88), 错误预警率为 81.82% (72/88)。暴发探测时间在 0~22 d 之间, 平均 6.25 d, 中位数为 4 d。预警提前时间在 -12~41 d 之间, 平均 9.75 d, 中位数为 5.5 d。被确认的 16 条时空重排扫描模型预警信号和对应的聚集性疫情信息, 见表 2。

表 2 被确认的 16 条预警信号和对应的流行性腮腺炎聚集性疫情信息

事件 编号	发生 日起	报告 日期	预警 日期	暴发探 测时间	预警提 前时间	预警信号包 括的街道数 <sup>a</sup>	RR 值	GLR 值	P 值
1	02-29	03-12	03-06	6	6	2	2.75	3.420	0.021
2	03-08	03-12	03-08	0	4	2	2.57	4.065	0.004
3	02-22	03-31	03-15	22	16	7	1.71	3.434	0.036
4	04-16	04-19	05-01	15	-12	1	1.87	4.594	0.002
5	03-23	04-20	04-04	11	16	2	1.76	3.018	0.061
6	05-15	05-30	05-25	10	5	3	3.04	4.428	0.008
7	06-05	06-07	06-09	4	-2	2	1.99	3.465	0.037
8	05-02	06-12	05-08	6	35	6	2.29	4.571	0.0016
9	05-01	06-15	05-05	4	41	2	1.72	3.085	0.048
10	06-18	06-19	06-29	11	-10	3	2.93	3.759	0.043
11	06-05	06-25	06-06	1	19	1	3.37	3.594	0.030
12	06-26	06-26	06-29	3	-3	3	2.93	3.759	0.043
13	10-08	10-08	10-08	0	0	3	3.40	4.181	0.015
14	10-11	10-18	10-15	4	3	1	3.07	3.152	0.046
15	10-16	11-16	10-18	2	29	6	2.10	3.574	0.010
16	12-16	01-05 <sup>b</sup>	12-17	1	9	1	2.31	3.589	0.016

注: a. 每宗疫情对应预警信号包括的街道数。b. 表格中除该项为 2013 年 1 月 5 日外, 其他时间均为 2012 年。

2.3 时空重排扫描模型和 CIDARS 模型比较 时空重排扫描模型产生的预警信号数为 88 次, 比 CIDARS 时间序列模型和时空序列模型预警信号数分别下降 72.41% 和 57.69%。时空重排扫描模型灵敏度高于 CIDARS 中时间序列模型 ( $\chi^2 = 15.789, P < 0.001$ ) 和时空序列模型 ( $\chi^2 = 13.228, P < 0.001$ )。时空重排扫描模型的信号阳性率显著高于 CIDARS 时间序列 ( $\chi^2 = 28.846, P < 0.001$ ) 和时空序列模型 ( $\chi^2 = 50.285, P < 0.001$ ), 见表 1。

3 讨 论

近 20 年来, 传染病预警方法研究迅速发展, 根据疾病流行的时间、空间、人群特征等属性, 主要分为时间模型、空间模型、时空模型和基于时空的多维模型等<sup>[15-17]</sup>。近年来, 中国疾病预防控制中心对 CIDARS 不断调整完善, 目前时间序列模型针对不同病种采用移动百分位数法、累积和控制图法、聚集性疫情法 3 种分类预警<sup>[5]</sup>, CIDARS 中时间模型是在县区的范围内进行运算, 深圳市人口密度极高, 丙类传染病发病率维持在较高水平, 不同社区人口密度差异较大, 造成各社区发病率差异也会较大, 规模较小的聚集性疫情可能难以被挖掘出, 同时受疾病周期性影响, 和历史数据比较时, 发病低谷年份的聚集性疫情难以辨别, 发病高峰年份又容易产生大量无用信号。假如在街道和社区层面运算, 由于传染病传播不会受行政区划的影响, 又会损失跨街道和社区的聚集性疫情的挖掘。

CIDARS 中时空序列模型是移动百分位数法与 kulldorff 空间扫描统计量法进行组合, 在县区范围先采用移动百分位数法探测, 存在异常变化的用空间探测模型扫描预警<sup>[2]</sup>, 该方法较时间序列模型效果有所提升, 但依然存在假信号较多等问题, 同时它和 kulldorff 时空模型的两个维度动态变化扫描的原理运算上有较大差异。时空重排扫描模型在时间和空间两个维度上同时不断变化扫描, 比较时空二维圆柱内外病例差异来判断是否存在聚集, 能够较为精确的在时空上进行定位, 同时可以较好的解决跨县区、跨街道疫情的挖掘, 在二个维度同时定位的情况下, 方便工作人员开展疫情核实。

本研究将 2012 年深圳市流行性腮腺炎时空重排扫描模型模拟探测结果和 CIDARS 中的时间序列模型 (移动百分位数法) 以及时空序列模型的探测结果进行对比, 发现时空重排扫描法预警信号数较少, 灵敏度较高, 错误预警率较低, 说明时空重排扫描法预警模型较 CIDARS 现行方法能更高效发现流行性腮腺炎聚集性疫情, 并在实际工作中可减轻工作人员的疫情核实工作, 研究结果和殷菲等<sup>[8]</sup>、孙玮璇等<sup>[18]</sup> 研究结果相一致。研究结果表明暴发探测时间中位数和平均数都比流行性腮腺炎平均潜伏期短, 说明在聚集性疫情发病高峰来临前, 模型就能做出早期预警。预警提前时间平均比疫情报告时间早了大约 10 d, 亦能说明模型能更早期发现聚集性疫情; 预警提前时间中有部分为负数, 可能和部分疫情病例数还很少时就报告有关。研究结果表明扫描预警信号包含的空间范围和聚集性疫情发生地吻合度较高, 基本都是在聚集性疫情发生地的周边街道, 这与实际情况很符合, 深圳系统上报的

流行性腮腺炎聚集性疫情基本都是发生在学校和托幼机构,这些学生基本都来自周边街道和社区范围。同时,本次研究发现,时空重排扫描模型的灵敏度不足 60%,信号阳性率不足 20%,错误预警率达 80% 以上,可能和空间扫描尺度还是过大有关<sup>[5]</sup>。流行性腮腺炎聚集性疫情主要发生在中小学校和托幼机构,聚集范围主要在社区甚至小区层面,研究是以街道为扫描尺度可能会影响结果,如果能够在社区范围或者学区范围开展研究,可能效果更好,也可能与本次研究只取了  $P < 0.05$  的预警信号有关。

本研究存在局限性包括,研究将深圳系统上报的聚集性疫情作为金标准来计算相关指标,虽然比现有研究<sup>[19]</sup>使用国家突发公共卫生事件的数据多,但疫情宗数仍可能比实际数少,依然有可能高估预警模型的灵敏度和错误预警率,现场调查时确定的聚集性疫情发生日期等信息也可能不够准确,会导致暴发探测时间和预警提前时间计算存在误差。本研究只取了  $P < 0.05$  的预警信号,可能会降低研究模型的敏感度和信号阳性率。本次研究过程中还发现 CIDARS 中存在问题,例如,深圳市近年来县区及街道有较大变化,新增和变更的街道病例数据在预警时和历史数据中街道编码不能对应,造成这些区域系统预警数据不可靠;预警系统中以县区为范围的预警难以发现不同县区相邻街道的暴发;时间模型参数设置未考虑不同传染病的潜伏期不同,过短的时间设置造成预警信号难以发出,过长的时间设置造成预警信号生成量太大;全国各地区人口密度、行政区域面积等差异较大,预警建模时应考虑到这些因素以提高预警效率。

#### 参考文献

- [1] 中国疾病预防控制中心. 全国传染病自动预警(时间模型)试运行工作方案[M]. 北京:中国疾病预防控制中心,2008,3-4.
- [2] 赖圣杰,廖一兰,张洪龙,等. 2011—2013 年国家传染病自动预警系统中时间模型和时空模型应用效果比较[J]. 中华预防医学杂志,2014,48(4):259-264.
- [3] 张洪龙,孙乔,赖圣杰,等. 移动百分位数法分地区设定预警阈值对传染病预警效果的影响分析[J]. 中华预防医学杂志,2014,48(4):265-269.

- [4] 陈碧云,高立冬,陈长,等. 我国传染病预警研究及工作现状[J]. 实用预防医学,2014,21(12):1537-1540.
- [5] 张洪龙,曾令佳,赖圣杰,等. 2016 年国家传染病自动预警信息系统运行情况分析[J]. 疾病监测,2018,33(2):159-167.
- [6] 方艳,宋铁,李灵辉,等. 广东省传染病自动预警系统运行现状分析[J]. 中华流行病学杂志,2013,34(8):800-803.
- [7] 吕炜,赖圣杰,唐忠,等. 广西壮族自治区 2009—2011 年传染病自动预警系统运行效果分析[J]. 中华流行病学杂志,2013,34(6):589-593.
- [8] 殷菲. 时空扫描统计量在传染病早期预警中的应用研究[D]. 成都:四川大学,2007.
- [9] Kulldorff M. A spatial scan statistic[J]. Commun Stat Theory Methods, 1997, 26(6):1481-1496.
- [10] Kulldorff M, Heffernan R, Hartman J, et al. A space-time permutation scan statistic for disease outbreak [J]. PLoS Med, 2005,2(3):216-224.
- [11] 赖圣杰,李中杰,金连梅,等. 传染病暴发早期预警系统评价内容及其指标[J]. 中华流行病学杂志,2009,30(6):637-640.
- [12] Wang X, Zeng D, Seale H, et al. Comparing early outbreak detection algorithms based on their optimized parameter values[J]. J Biomed Inform, 2010,43(1):97-103.
- [13] Watkins RE, Eagleson S, Veenendaal B, et al. Applying cusum based methods for the detection of outbreaks of Ross River virus disease in Western Australia[J]. BMC Med Inform Decis Mak, 2008,8:37.
- [14] Hutwagner L, Browne T, Seaman GM, et al. Comparing aberration detection methods with simulated data[J]. Emerg Infect Dis, 2005,11(2):314-316.
- [15] 刘天,姚梦雷,陈红缨,等. 单纯 ARIMA 模型和基于季节性分解的 ARIMA 模型在丙肝发病率中的预测效果比较[J]. 实用预防医学, 2019,26(3):278-281.
- [16] 焦锋,董蓬玉,刘晓强,等. 传染病预警方法及应用概述[J]. 中国社会医学杂志,2018,35(4):340-343.
- [17] 朱婷婷,孙利文,常姗姗,等. 比数图法在北京市怀柔区重点传染病预警中的方法探讨[J]. 实用预防医学, 2017,24(10):1270-1273.
- [18] 孙玮璇,廖志林,毛绍菊,等. 基于前瞻性时空重排扫描统计量的手足口病早期预警方法及其实证研究[J]. 南昌大学学报(医学版),2015,55(6):80-83.
- [19] 孙秋云. 基于网络直报下的传染病自动预警信息系统运行效果分析[J]. 中国卫生统计,2017,34(3):465-466.

收稿日期:2019-04-25