

信号强度指数在传染病自动预警中的应用及评价

周鼎伦 杨维中 孙乔 赖圣杰 张洪龙 李中杰 吕炜 兰亚佳

【摘要】目的 为改进国家传染病自动预警系统(CIDARS),研究引入预警信号强度指数(SSI),以评估 SSI 在改善预警信号灵敏度中的作用。**方法** 采用随机数字表法抽取全国 2007—2011 年痢疾病例报告数据和 2010—2011 年的预警信号数据,并计算 SSI 和疑似事件预警信号关联比(ER)。分析 SSI 与 ER 的相关关系,并用多因素 logistic 回归分析法探索 SSI 对 ER 的影响。**结果** 2010—2011 年 CIDARS 共发出的 9 620 个预警信号中,传染病暴发疑似事件信号数 74 条(0.77%)。SSI 分级与 ER 之间有良好的相关性:疑似事件相关联的信号 SSI 中位数为 4.0,远高于非疑似事件相关联的信号(1.7);随着 SSI 的增加,事件关联比增加($r=0.917$),当 SSI 超过 20 时,事件关联比可以高达 20;多因素 logistic 回归模型分析发现,SSI 与疑似事件信号的 $OR(95\%CI)$ 值为 2.52(2.04~3.12);流行季 SSI 对疑似事件概率的指示作用高于非流行季。**结论** SSI 与 ER 的相关性很好,尤其是对于暴发规模相对较大的传染病暴发事件,且流行季的作用较非流行季明显。

【关键词】 传染病; 疾病爆发流行; 信号强度

基金项目: 卫生部卫生公益性行业科研专项(201202006); 国家科技重大专项(2012ZX10004-201)

Application and evaluation of signal strength indicator in communicable disease automatic early warning system Zhou Dinglun*, Yang Weizhong, Sun Qiao, Lai Shengjie, Zhang Honglong, Li Zhongjie, Lyu Wei, Lan Yajia.*West China of Public Health, Sichuan University, Chengdu 610041, China
Corresponding author: Lan Yajia, Email: lanyajia@sina.com

【Abstract】Objective To explore the effect of signal strength indicator (SSI) in improving sensitivity of China Infectious Diseases Automated-alert and Response System (CIDARS). **Methods** Diarrhea cases in 2007–2011 and early warning signals in 2010–2011 were selected by using random digital table method. Then, SSI and event-related ratio (ER) were calculated. The relationship between ER and SSI was analyzed, and the effect of SSI on ER was explored by using multiple logistic regression analysis. **Results** 9 620 early warning signals in 2010–2011 were generated in two years. Of these, 74, or 0.77% were defined as suspected outbreak signal. The median of SSI related with suspected outbreak signal was 4.0, which was much higher than non-suspected outbreak signal (1.7). ER was significantly correlated with SSI ($r=0.917$). SSI classification has a good correlation between the ER, ER exceeded 20 after SSI reached 20. The multivariate logistic regression analysis showed OR of SSI related with suspected outbreak signal was 2.52 (95%CI 2.04–3.12). Compared with non-epidemic season, the relationship of SSI and ER in epidemic season was much higher. **Conclusion** SSI was closely related with ER. The relationship was much closer in large scale outbreak and epidemic season, and compared to non-epidemic, the effect of epidemic season is more obvious.

【Key words】 Communicable diseases; Disease outbreaks; Signal strength indicator

Fund program: Research Projects of Public Welfare Industry in Public Health Ministry (201202006); National Science and Technology Major Project of China(2012ZX10004-201)

传染病暴发早期预警对早期识别传染病暴发,

并迅速采取有效的控制措施,降低相关传染病的发病率和死亡率,具有重要的公共卫生意义^[1-6]。我国于 2008 年 4 月启动了国家传染病自动预警系统(China Infectious Diseases Automated-alert and Response System, CIDARS)。CIDARS 在国家传染病防控工作中发挥着重要作用,同时也存在一些问题,比如预警阳性率低,初步判断为疑似事件的信号数约占发出信号总数的 1%^[7-9];预警信号的发出

DOI: 10.3760/cma.j.issn.0253-9624.2016.02.016

作者单位: 610041 成都, 四川大学华西公共卫生学院(周鼎伦、兰亚佳); 中国疾病预防控制中心(杨维中); 中国疾病预防控制中心传染病预防控制处, 传染病监测预警中国疾病预防控制中心重点实验室(赖圣杰、张洪龙、李中杰); 上海市浦东新区疾病预防控制中心(孙乔); 广西壮族自治区疾病预防控制中心(吕炜)

通信作者: 兰亚佳, Email: lanyajia@sina.com

与否只反映了当前期传染病病例报告数是否超过了既往同期传染病报告数阈值,是一种定性预警,如能在此基础上提供预警信号强度信息,对预警响应工作的指导意义将更强。因此,本研究通过引入预警信号强度指标,评估该指标在改善预警信号灵敏度中的作用,对提高 CIDARS 系统的有效性,增强预警响应行动的针对性具有现实意义。

资料与方法

1. 数据来源:以痢疾为目标疾病,所有数据及分析均是以全国法定报告传染病监测数据和 CIDARS 中痢疾信息为基础。纳入了 2007—2011 年全国的痢疾病例报告数,以及 2010—2011 年的痢疾预警信号数。因为 CIDARS 预警系统需要 3 年的数据作为基线。故 2010 年的预警信号强度是以 2007—2009 年病例报告数为基线,2011 年的预警信号强度是基于 2008—2010 年的基线数据计算而来的。由于数据库记录庞大,预警信号强度分析没有采用全部样本,而是采取抽样的方法,对样本数据进行分析。从 2010—2011 年全国 CIDARS 的预警信号库中,先抽取有痢疾疑似事件报告的全部县,共 53 个县;再从没有疑似事件的县中按 12% 的比例用随机数字表法抽取 300 个县,合计抽取 353 个县。其中 1 个县没有病例报告,分析时剔除该县。最终,纳入分析的是 352 个县(53 个有疑似事件的县,299 个只有预警信号的县)。数据抽取框架见表 1。研究样本两年共发出预警信号 9 620 个,病例报告数为 222 816 例。

2. 信号强度指数(signal strength indicator, SSI): CIDARS 预警算法采用的是移动百分位数法^[10]。该方法以 7 d 为一个观察期数据块,将当前期数据块与既往 3 年同期及前后各摆动 2 个数据块(共 15 个数据块)相比较,若当前期病例数超过基线数据的某百分位数值对应的病例数(预警阈值),系统即生成预警信号。因此,理论上该方法除了可以定性比较当前期病例观察数与预警阈值大小外,还可以提供当前期病例观察数与预警阈值之比。因此,本研

究构建了 SSI 来表达预警信号的强度。 $SSI = C_i / E_i$,其中 C_i 为预警信号对应的当日病例数, E_i 为预期病例数,即前 3 年同日加上前后各摆动 3 d 形成的数据块的平均病例数。SSI 越大,表示信号越强。

3. 事件关联比(event-related ratio, ER):为了便于描述,本研究将与疑似事件有联系的预警信号数与总预警信号数的百分比称为 ER。 $ER = \text{Signal}_e \times 100 / \text{Signal}_t$,其中,Signal_e 为与疑似事件有联系的预警信号数,Signal_t 为总预警信号数。

4. 统计学分析:所有统计分析均使用 R 3.0.1 软件完成。根据 CIDARS 的运行过程可看出,疑似事件是 CIDARS 的一个重要指标,信号灵敏度的计算是根据疑似事件获得,而是否需要进一步做出现场处理,也是依据疑似事件做出判断。为此,将判断为疑似事件的预警信号数与预警信号总数的综合信息 ER 作为衡量指标,来统计预警 SSI 在传染病预警中的作用。

在用列联表法分析不同强度等级 SSI 与 ER 关系的基础上,纳入流行季节、流行水平、预警界值,采用多因素 logistic 回归模型分析影响 ER 的因素。流行季节是根据 2010、2011 两年逐月报告病例数确定的。考虑到各省的季节性有差别,故在确定是否处于流行季节时,以省为单位。根据按月报告病例数的分布特征,如果某个省的月报告病例占该省同年报告总数的比例连续超过 9.5%,则涉及的月份被判断为流行季节,否则为非流行季节。流行水平是根据 2010 和 2011 两年的年平均报告发病率,以发病率 1/10 万、5/10 万、10/10 万为分界点,从低到高分 4 个等级。痢疾流行水平属于低、中、较高和高者分别占 35.3%、52.3%、6.9% 和 5.5%。

结 果

1. 描述信号的特征:发出的 9 620 个预警信号中,有 74 条判断为传染病暴发的疑似事件,占 0.77%,而超过 99% 的预警信号未关联疑似事件,可以视为假阳性信号。统计不同月份的预警信号的 ER,结果见表 2。可见不同月份的 ER 不同,但未表

表 1 研究数据抽样框架说明

数据库	原数据库		样本数据库		抽样说明
	记录数	覆盖范围	记录数	覆盖范围	
预警信号库	67 112	2010—2011 年产生的预警信号,共覆盖 2 556 个县	9 620	352 个县	2010—2011 年有疑似事件的县全抽,没有疑似事件的县按 12% 随机抽取
病例报告库	1 456 049	2007—2011 年传染病直报系统报告的病例,共覆盖 3 116 个县	222 816	352 个县	包括预警信号样本库覆盖县的 5 年报告病例数据

现出明显的规律。

表 2 不同月份预警信号的疑似事件关联比

月份	预警信号数	疑似事件信号	事件关联比
1	400	4	1.00
2	369	0	0.00
3	561	1	0.18
4	669	1	0.15
5	909	8	0.88
6	1 021	11	1.08
7	1 226	8	0.65
8	1 275	9	0.71
9	1 225	25	2.04
10	829	5	0.60
11	656	2	0.30
12	480	0	0.00

2. 预警信号强度与疑似事件的关系:所有发出的预警信号中,其 SSI 中位数为 1.7, P_{25} 和 P_{75} 分别为 1.2 和 2.6。区分为疑似事件信号和非疑似事件信号后,可看出疑似事件相关联的信号 SSI 中位数为 4.0,远高于非疑似事件相关联的信号(表 3),提示 SSI 的大小与疑似事件有很强的联系。

进一步分析 SSI 与疑似事件的关系,结果见表 4。随着预警 SSI 的增加,ER 增加, r 值为 0.917 ($P < 0.001$)。当 SSI 超过 20 时,ER 可以高达 20,显示高强度 SSI 与 ER 具有良好的相关性。

多因素 logistic 回归分析发现,SSI、流行季节是影响 ER 的因素,结果有统计学意义。其中,SSI 的 OR 值为 2.52,表示信号强度每增加一个等级,信号与疑似事件关联的概率增加 2.52 倍。同理,流行季节发出的预警信号,信号与疑似事件关联的概率比非流行季节预警信号高 4.86 倍。而流行水平、预警界值两个因素的影响未达到显著性水平。详见表 5。

区分流行季与非流行季,观察 SSI 与疑似事件概率的关系,结果见图 1。可看出,无论是流行季还是非流行季,疑似事件概率均随着 SSI 的增加而增加,且流行季增加幅度高于非流行季。然而,需要注

意的是,在 SSI 低于 20 时,增加幅度较缓,而 20~60 之间,增加幅度很大。提示预警信号强度较小时,SSI 对疑似事件的判断帮助有限,而当预警信号强度较大(>20),SSI 对疑似事件的判断效率很高。

表 4 不同等级信号强度指数的事件关联比

信号强度指数	预警信号数	疑似事件信号	事件关联比
0~	891	3	0.34
1~	4 834	20	0.41
2~	1 878	9	0.48
3~	782	4	0.51
4~	360	8	2.22
5~	518	7	1.35
10~	131	8	6.11
15~	166	3	1.81
20~	60	12	20.00
合计	9 620	74	0.77

表 5 疑似事件信号影响因素的 logistic 回归分析

因素	OR 值	95% CI 值		P 值
		下限	上限	
信号强度指数	2.52	2.04	3.12	<0.001
流行季节	4.86	2.68	8.83	<0.001
流行水平	0.98	0.72	1.32	0.892
预警界值	1.02	0.99	1.05	0.250

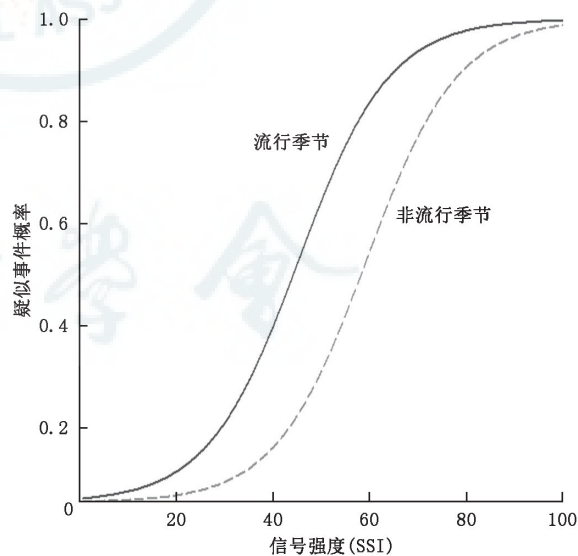


图 1 信号强度指数与疑似事件概率的关系

表 3 不同信号类型的信号强度指数

信号类型	n	均数	标准差	最小值	最大值	中位数	P_{25}	P_{75}
所有信号	9 620	2.7	6.6	0.1	510.0	1.7	1.2	2.6
疑似事件信号	74	18.9	63.8	0.4	510.0	4.0	1.6	13.0
非疑似事件信号	9 546	2.6	3.3	0.1	125.0	1.7	1.2	2.6

讨 论

痢疾是我国一种重要的乙类传染病,年报告发病率及预警信号发出数均位居前列^[7-9]。本研究随机抽取的县中,总信号数为 9 620 个,而疑似 ER 仅为 0.77%,与国内一些研究结果类似^[7-9]。引入 SSI 指标后,疑似事件相关联的 SSI(18.9)远高于非疑似事件信号的 SSI(2.6),提示 SSI 对推断预警信号是否为疑似事件具有重要作用。进一步分析发现,随着 SSI 强度的增加,与疑似事件的关联性就会增大。当 SSI 超过 20,20% 的信号被判断为疑似事件信号。因此,如能在发出预警信号的同时,告知预警信号强度,对基层传染病防控工作判断该预警信号的价值有很大的帮助。

分析影响疑似事件判断的因素发现,除 SSI 外,流行季节也是一个重要的因素。从疑似事件概率与 SSI 关系图可见,同等数值的 SSI,在流行季与疑似事件的关联概率高于非流行季。可能与非流行季的病例数少,故相比较而言变异度更大有关。

本研究也存在一些不足之处。首先是疑似事件例数很少,9 620 个预警信号中只有 74 个信号被判断为疑似事件。过少的疑似事件不利于深入、综合地探讨 SSI 与其他各种因素(如流行季、发病水平、地域等)的关系,其结果的稳定性也受到影响。其次,从表 4 可看出,当 SSI 值很小时,也有可能被判断为疑似事件,原因有待进一步分析查明。一个可能的原因是移动平均数法是以县为单位进行运算的,而有些暴发事件发生在学校等小区域,事件病例数少、规模较小,防控人员根据其他渠道(如学校报告)获取信息,继而在响应系统中将其标记为疑似事件^[11-12]。因此,发现传染病暴发的途径是多种多样的,不应期待预警系统能完全替代其他的暴发发现方式^[13-14]。再次,SSI 与疑似事件概率的关系图可看出,SSI 数值高于 20 之后,关联性快速增加,而低于 20 则上升速度较缓慢,因此对于局部和暴发规模不大的传染病暴发事件的贡献不明显。

综上,预警信号强度指标能起到提高传染病预警效能的作用,尤其是对于暴发规模相对较大的暴发事件,且在流行季的作用较非流行季明显。可以考虑在

CIDARS 中增加预警信号强度指标,更好地指导基层传染病预警工作。但受限于疑似事件样本数较小等因素,尚需继续开展深入研究,以进一步评估预警强度指标在传染病预警工作中的作用及分析影响因素。

参 考 文 献

- [1] Hutwagner L, Thompson W, Seeman GM, et al. The bioterrorism preparedness and response Early Aberration Reporting System (EARS)[J]. *J Urban Health*, 2003, 80(2 Suppl 1): i89-96.
- [2] Tsui FC, Espino JU, Dato VM, et al. Technical description of RODS: a real-time public health surveillance system[J]. *J Am Med Inform Assoc*, 2003, 10(5): 399-408. DOI: 10.1197/jamia.M1345.
- [3] 李中杰,马家奇,赖圣杰,等.2011—2013 年国家传染病自动预警系统运行结果分析[J]. *中华预防医学杂志*, 2014, 48(4): 252-258. DOI: 10.3760/cma.j.issn.0253-9624.2014.04.004.
- [4] Lombardo J, Burkom H, Elbert E, et al. A systems overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II) [J]. *J Urban Health*, 2003, 80(2 Suppl 1): i32-42. DOI: 10.1089/153871303771861441.
- [5] Mandl KD, Overhage JM, Wagner MM, et al. Implementing syndromic surveillance: a practical guide informed by the early experience[J]. *J Am Med Inform Assoc*, 2004, 11(2): 141-150. DOI: 10.1197/jamia.M1356.
- [6] Nkuchia M, M'ikantha, Lynfield R, et al. *Infectious Disease Surveillance*[M]. London: Blackwell, 2007: 349-358.
- [7] 刘润友,杨长虹,罗湘蜀,等.2008—2013 年四川省传染病自动预警系统应用效果分析[J]. *现代预防医学*, 2014, 41(21): 3918-3921.3978.
- [8] 鲁琴宝,徐旭卿,林君芬,等.2009—2012 年浙江省传染病自动预警系统预警结果分析[J]. *中国预防医学杂志*, 2014, 15(7): 654-658.
- [9] 陈碧云,高立冬,李忠杰,等.湖南省传染病自动预警系统预警效果分析[J]. *实用预防医学*, 2012, 19(12): 1767-1771. DOI: 10.3969/j.issn.1006-3110.2012.12.002.
- [10] 杨维中. *传染病预警理论与实践*[M]. 北京:人民卫生出版社, 2012: 166-167.
- [11] 吴昊澄,林君芬,何凡,等.浙江省 2005—2012 年学校突发公共卫生事件流行病学分析[J]. *中国学校卫生*, 2013, 34(9): 1114-1116.
- [12] 尤恩情,吴金菊,赵科伙,等.传染病自动预警系统运行状况及预警效果评价[J]. *安徽预防医学杂志*, 2014, 20(5): 323-326.
- [13] 杨维中,兰亚佳,李中杰,等.国家传染病自动预警系统的设计与应用[J]. *中华流行病学杂志*, 2010, 31(11): 1240-1244. DOI: 10.3760/cma.j.issn.0254-6450.2010.11.009.
- [14] 魏晶,王海涛,刘小晖,等.青岛市学校因病缺课症状监测直报系统准确性评估[J]. *中国学校卫生*, 2012, 33(9): 1123-1126.

(收稿日期:2015-03-04)

(本文编辑:吕相征)