

上海世界博览会期间浦东新区呼吸道症状监测数据特征与稳定性分析

王小皙 杨维中 孙乔 李中杰 周鼎伦 叶楚楚 兰亚佳

【摘要】 目的 分析上海世界博览会(世博会)浦东新区发热呼吸道症候群监测数据,揭示监测数据特征及监测系统运行的稳定性。方法 分析数据源来自 2010 年 5 月 1 日至 10 月 31 日期间浦东症状监测系统内的发热呼吸道症候群监测数据,以流感样病例(ILI)监测数据作为参照,在数据特征描述的基础上,进行两者相关性分析,并通过比较症状监测前后期的数据变异性与上报滞后时间来评价数据质量及稳定性。结果 监测显示,发热呼吸道症候群监测序列于 9 月中下旬达到高峰,具有周末效应及节假日效应,与 ILI 监测数据同一天的相关性最佳($r=0.596, P<0.05$)。监测前期(2010 年 5 月 1 日至 7 月 31 日),发热呼吸道症候群与 ILI 监测序列没有表现出明显的相关性($r=-0.058, P>0.05$);而在监测后期(2010 年 8 月 1 日至 10 月 31 日),两者时间趋势一致,同一天的 r 值达到 0.798($P<0.05$)。分析数据的变异性发现监测后期数据变异性小于前期,同时监测后期的数据上报质量高于前期。综合与参照序列的相关性、数据变异性及报告质量分析,提示症状监测系统后期的数据稳定性高于前期。结论 发热呼吸道症候群监测数据在监测系统运行一段时间后才会趋于稳定,在设计针对大型活动的监测系统时应该预留足够的试运行期,并在选用短基线预警模型时应考虑监测数据存在周末与节假日效应。

【关键词】 症状监测;世界博览会;稳定性

Characteristics and stability of surveillance data on respiratory syndrome, during the Shanghai World Expo in Pudong New District WANG Xiao-xi¹, YANG Wei-zhong², SUN Qiao³, LI Zhong-jie², ZHOU Ding-lun¹, YE Chu-chu³, LAN Ya-jia¹. 1 West China School of Public Health, Sichuan University, Chengdu 610041, China; 2 Chinese Center for Disease Control and Prevention; 3 Shanghai Pudong New District Center for Disease Control and Prevention

Corresponding author: LAN Ya-jia, Email: lanyajia@sina.com

This work was supported by a grant from the National Science and Technology Major Project of China (No. 2009ZX10004-201).

【Abstract】 Objective To reveal the characteristics and stability of the system through the analyzing the surveillance data of respiratory-feverous syndrome via the syndromic surveillance system which was established during the Shanghai World Expo in Pudong New District and provide references for the development and operation optimization on this Mass Gatherings Surveillance Systems. Methods Data used was from the surveillance data of respiratory-feverous syndrome collected from Pudong New District Syndromic Surveillance System, through May 1 to October 31, 2010. On the basis of description of data characteristics, correlation analyses were conducted, when compared to the surveillance data of respiratory-feverous syndrome and Pudong influenza-like illness (ILI) used as reference. Comparison of variances on the surveillance data and the report lag time of the earlier and later surveillance periods were also carried out to evaluate the quality and stability of data. Results Reports on the respiratory-feverous syndrome showed a peak in late September with day-of-week effects and holiday effects. Correlation between respiratory-feverous syndrome and ILI was the strongest in the same day ($r=0.596, P<0.05$). In the earlier surveillance period from 2010-05-01 to 2010-07-31, the correlation between respiratory-feverous syndrome and ILI was not obvious ($r=-0.058, P>0.05$); however, the two-time series showed consistent trend with the

DOI: 10.3760/cma.j.issn.0254-6450.2012.06.004

基金项目:国家科技重大专项(2009ZX10004-201)

作者单位:610041 成都,四川大学华西公共卫生学院劳动卫生教研室(王小皙、周鼎伦、兰亚佳);中国疾病预防控制中心(杨维中、李中杰);上海市浦东新区疾病预防控制中心(孙乔、叶楚楚)

通信作者:兰亚佳, Email: lanyajia@sina.com

correlation coefficient as 0.798 ($P < 0.05$), in the later period from 2010-08-01 to 2010-10-31. In addition, variability of the surveillance data on respiratory-feverous syndrome was less in the later period than in the earlier one, with quality of the report on related data better in the later period. Analyses on the correlations of reference sequence, variability and quality of report indicated that the stability of the later surveillance period was better than the earlier one. **Conclusion** Only with the operation of syndromic surveillance system for a certain period of time, could data in the system maintain stability. Surveillance data showed both day-of-week effects and holiday effects, suggesting that there was a need to choose early warning models with short baseline data.

【Key words】 Syndromic surveillance; World expo; Stability

为了保障大型活动期间的公共卫生安全,不少国际大型活动都启动了症状监测系统来加强常规病例报告监测^[1-3]。2010年5—10月上海世界博览会(世博会)期间,浦东新区在世博园周边建立了症状监测预警系统。由于为大型活动建立的症状监测系统具有临时性,缺乏历史基线数据,需要建立新的数据收集渠道。分析新建渠道所收集信息的数据特征及稳定性,可以从一定角度反映症状监测系统的运行状况,对今后建立类似症状监测系统具有参考意义。本研究通过描述发热呼吸道症候群监测数据的特征,与流感样病例(ILI)监测数据进行相关性分析以及数据变异性和报告质量分析,以评估数据的稳定性,估计大型活动建立症状监测系统所需的测试期。

资料与方法

1. 分析数据源:选择浦东症状监测系统中的发热呼吸道症候群监测数据,选择这一症候群的原因是它对应的目标疾病众多,包括流感、SARS、人禽流感、军团菌病等10种传染病,此外,国内外的症状监测系统中多数将此症候群列为最主要的监测目标^[4,5]。分析数据源主要来自浦东新区34家监测医院的信息管理系统(HIS系统)。每日24时各医院的数据通过前置机自动上传至浦东新区卫生信息平台。收集数据时间从2010年5月1日至10月31日,共184 d。发热呼吸道症候群定义为有咽痛、咳嗽、胸痛、咯血、呼吸困难症状之一,且伴有发热。目标症候群的定义通过文献法、德尔菲法、历史数据分析法确定。

2. 参照数据源:流感是发热呼吸道症候群中主要的目标疾病,ILI监测可以指示流感的流行。浦东ILI监测系统是根据全国流感监测方案建立的常规监测系统之一,已经稳定运行10年,选择ILI监测数据作为参照数据源来评价症状监测系统数据稳定性是相对合理的,若与ILI监测数据的相关性高,则提示发热呼吸道症候群监测数据的稳定性好。根据全国流感监测方案要求,浦东新区建立的ILI监测哨点

医院共20家,包括国家级流感监测点2个,市级流感监测2个以及区级流感监测点16个,与症状监测医院共有13家医院重叠。各哨点医院监测诊室医务人员在诊疗中发现符合ILI定义的患者做好专册登记。负责流感监测的专职人员每日统计前一日各监测诊室门急诊ILI数,上报到中国流感监测信息系统。收集与分析数据源同期(即2010年5月1日至10月31日)的ILI监测数据。ILI的定义为发热(体温 $\geq 38^\circ\text{C}$),伴咳嗽或咽痛者。

3. 统计学分析:描述发热呼吸道症候群监测数据的特征,包括报告症状的比例,56 333例患者的人口统计学特征分布,时间序列特征以及周末效应和节假日效应。在此基础上,采用7 d移动平均方法对发热呼吸道症候群与ILI监测序列做预处理^[6],消除周末效应,以更好的反映其时间趋势。对预处理后的发热呼吸道症候群和ILI监测数据做时间交叉相关分析^[4,7]。相关性分析可以得到发热呼吸道症候群对应ILI监测数据前后2周的相关系数,并且确定两者相关性最高时的时滞关系。在数据稳定性分析中,还采用一阶差分($d_t = x_t - x_{t-1}$)的方差值来描述发热呼吸道症候群监测数据波动性的变异程度,以及选用上报滞后时间来评价数据的报告质量。采用R 2.10.0软件进行图形生成和统计学分析。

结 果

1. 报告症状的比例:发热呼吸道症候群包含6个症状,分别为发热、咽痛、咳嗽、胸痛、咯血和呼吸困难症状。发热为监测症候群定义的必需症状,其他为伴随症状,表1显示了这些伴随症状的报告比例,咳嗽且伴发热为最常见的症状,占到68.33%,其次为咽痛且伴发热症状,为51.47%,患有严重症状如发热伴胸痛、咯血和呼吸困难等的患者数少。

表1显示,将监测对象按0~14岁、15~59岁和 ≥ 60 岁分为3个年龄组,描述了各年龄组中伴随症状的报告比例。在0~14岁及 ≥ 60 岁组中,出现频率最高的症状为咳嗽伴发热;而在15~59岁组中,出现频率最高的为咽痛伴发热。

表 1 发热呼吸道症状在各年龄组中的报告比例(%)

症状	年龄组(岁)			合计 (n=56 333)
	0~14 (n=27 065)	15~59 (n=21 198)	≥60 (n=8070)	
咳嗽	75.62	56.42	75.18	68.33
咽痛	32.67	74.44	54.19	51.47
胸痛	0.07	0.93	2.00	0.67
咯血	0.02	0.18	0.69	0.18
呼吸困难	0.06	0.35	1.62	0.40

2. 人口统计学特征:发热呼吸道症候群患者中男女比例为 1.22 : 1,以男性居多,门急诊 ILI 男女性比例为 1.17 : 1,两者性别比例相近。0~14 岁组发热呼吸道症候群就诊人数最多,占 48.04%,其次为 15~59 岁组;ILI 也集中在 0~14 岁组,占 ILI 监测人数的 69.99%(表 2)。

表 2 发热呼吸道症候群患者和 ILI 的一般人口学特征

人口学特征	发热呼吸道症候群		ILI	
	例数	构成比(%)	例数	构成比(%)
年龄组(岁)				
0~14	27 065	48.04	67 477	69.99
15~59	21 198	37.63	23 878	24.77
≥60	8 070	14.33	5 053	5.24
合计	56 333	100.00	96 408	100.00
性别				
男	30 957	54.95	52 026	53.96
女	25 376	45.05	44 382	46.04

3. 时间序列特征:采用时间序列分布图来描述发热呼吸道症候群监测数据的时序特征。图 1A 为发热呼吸道症候群与参照序列 ILI 监测数据的原始时间分布图,图 1B 为平滑处理后的时间分布图。2010 年 9 月上旬,发热呼吸道症候群报告人数开始增加,到 9 月中下旬达到高峰,10 月报告人数下降(图 1B)。如图 1B 所示,以 2010 年 8 月 1 日为分界点(监测分期线)将整个监测阶段划分为前期与后期。在监测前期 2010 年 5 月 1 日至 7 月 31 日,发热呼吸道症候群与 ILI 监测序列的时间趋势没有明显的相关性,而在后期 2010 年 8 月 1 日至 10 月 31 日症候群序列与参照序列的时间趋势一致。

4. 周末效应和节假日效应:发热呼吸道症候群数据序列表现出周末降低现象,而 ILI 监测不具有周末效应,监测数据呈随机波动(表 3)。观察期内主要节日为“十一”国庆节,以此来分析节假日效应,“节日周”定义为 10 月 1~7 日。比较节日周、节前一周和节后一周的监测数据来反映节假日效应。结果显示,发热呼吸道症候群和 ILI 在节前、节中和节后的平均报告数差异有统计学意义,逐周降低(表 3)。

表 3 周末效应和节假日效应

数据源	平均日报告数		P 值*	平均日报告数				P 值*
	平日	周末		节前周	节日周	节后周	节后周	
发热呼吸道症候群	312.33	291.30	0.194	486.43	427.14	353.14	0.000	
ILI	523.01	526.24	0.838	611.29	580.00	456.57	0.004	

注:*单因素方差分析

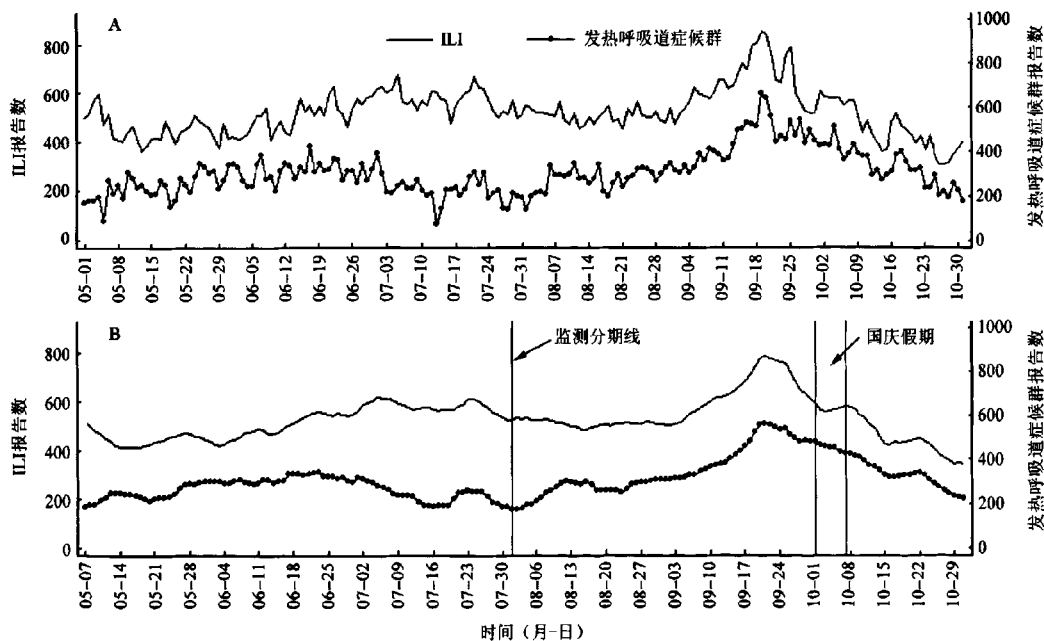


图 1 预处理前后发热呼吸道症候群与 ILI 时间序列

鉴于发热呼吸道症候群监测数据具有周末效应,运用7 d移动平均方法对该时间序列做预处理,消除周末效应。症候群和ILI监测数据表现出节假日效应,因此在分析时排除国庆节节假日周的数据,以消除节假日效应的影响。为了在相关分析中保证两种监测数据源的可比性,对两种监测数据都做了相同的序列预处理。预处理后的监测序列时间分布如图1B所示,图中第2、3垂直线之间为消除节假日效应需切除的序列段,在分析时将被排除。

5. 与参照序列的相关分析:对预处理后的发热呼吸道症候群和ILI监测数据做时间交叉相关分析,交叉相关系数见图2。以症候群监测数据提前天数为横坐标,两种数据源的相关系数为纵坐标作图。图2展示了症候群监测数据提前ILI监测数据-14~11 d的相关性,因为从对应ILI前14 d到后11 d内数据的相关系数具有统计学意义。图中可发现症候群监测数据与ILI监测数据同一天的相关系数达到最大值0.596($P < 0.05$)。开展症状监测与ILI监测的医院重叠比例较大,考虑数据源重叠对相关性的影响,分别对同时开展两种监测的医院,以及仅开展一种监测医院的两种数据源的相关性进行分析。开展两种监测的医院,症候群监测数据与ILI监测数据同一天的相关系数为0.565($P < 0.05$);而只开展症候群监测或ILI监测的医院,两种数据源同一天的相关系数为0.464($P < 0.05$)。

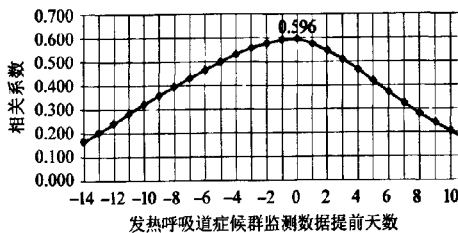


图2 时间交叉相关系数

根据时间序列特征分析(图1B)可见,症候群与ILI监测序列在8月1日前后两个监测阶段呈现出不同的相关性。在监测前期,发热呼吸道症候群与ILI监测序列同一天的相关系数为-0.058($P > 0.05$),没有明显的相关性;而在监测后期两者的时间趋势一致,同一天的相关系数达到0.798($P < 0.05$)。

6. 数据波动性与报告质量:通过分析数据的波动性与报告质量来描述数据的稳定性。通过比较症候群监测数据前后期一阶差分的变异程度即方差值,来分析数据的波动性。一阶差分值 $d_t = x_t - x_{t-1}$,表示当日与前一日发热呼吸道症候群报告数之差,

其分布情况如图3所示,监测前期一阶差分值的全距比后期大($R_{前期}=325, R_{后期}=279$)。对监测前后期一阶差分的方差值做单因素方差分析,前期的方差值为3704.37,后期为2755.93 ($F=1.786, P=0.183$)。提示发热呼吸道症候群监测前期的数据波动性大于后期。

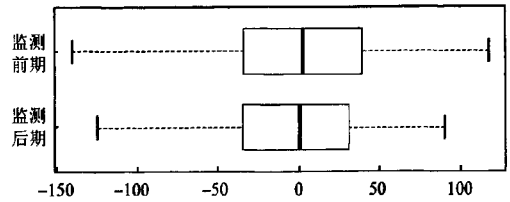


图3 监测前后期一阶差分箱形分布图

考虑到建立的症状监测系统是每天实时计算并预警,因此将滞后时间为2 d及以上的数据视为上报质量不好。滞后时间=上报时间-监测日期。其中,“上报时间”和“监测日期”均为系统自动生成的数据。其中,“上报时间”是数据传输时系统自动生成的日期,“监测日期”是患者看病时的时间。以滞后时间考核数据上报质量,发现多数数据的监测时间和上报时间相差在1 d以内,但也存在滞后时间较长的情况,其中滞后时间最长的是71 d,出现在监测前期。监测前期滞后2 d及以上数据比例占到病例记录总数的4.00%,监测后期仅有0.82%,差异有统计学意义($\chi^2=652.304, P < 0.001$),提示监测后期的数据上报质量高于前期。

讨论

国内外已有研究探讨多种数据源之间的关系^[7,8],但很少有研究评价症状监测系统数据的稳定性。数据的稳定性可以反映系统运行的稳定性,而系统稳定性又是评价监测系统的一个重要指标^[9]。本研究以2010年上海世博会期间的浦东新区发热呼吸道症候群报告数为分析数据,通过描述发热呼吸道症候群监测数据的特征,与ILI监测数据进行相关性分析以及数据变异性和报告质量分析,来评估数据的稳定性,进而反映症状监测系统的稳定性,为今后大规模人群聚集活动监测系统的建立及运行优化提供借鉴与参考。

专门针对大规模人群聚集活动建立的症状监测系统具有其必要性。这种监测系统以预警为目标,实现多种疾病早期症状的同步监测,具有临时性,并且收集医院、药店、学校、宾馆等多种监测数据,都是与常规监测系统不同之处,可以与常规监测系统相

互补充。在没有相应常规监测系统的地方特别需要专门建立。常规监测不是为特定活动的传染病预警而设计,所以在目标人群针对性、信息源范围、监测覆盖区域等达不到这类活动卫生安全保障的要求。

34家症状监测医院与20家ILI监测哨点医院共有13家重叠。症状监测医院未重叠的多为社区医院,而ILI监测哨点医院未重叠的包括国家级与市级监测医院。流感监测方案中针对儿童流感专门设置了国家级流感监测点,而新建的症状监测系统未对不同年龄人群进行重点监测,再加上新增数据收集的人员配置等问题,导致发热呼吸道症候群监测发现的患者数小于ILI监测。

选取的分析数据源与参照数据源的监测系统不同,包括监测的症候群定义不同,监测信息的收集方式不同等,因而两者的相关性不一定显著。发热呼吸道症候群与ILI监测数据的相关性大小,是评价症状监测数据稳定性的指标之一,全序列相关分析可见症状监测数据总体上稳定。进一步研究发现,监测后期发热呼吸道症候群与参照数据源的相关性明显高于监测前期,且通过比较发热呼吸道症候群监测前后期的数据变异性,发现监测后期的数据变异程度小于监测前期,同时监测后期的数据上报质量高于前期。

综合与参照序列的相关性、数据变异性及报告质量分析,提示症状监测系统后期的数据稳定性高于前期。发热呼吸道症候群监测数据在监测系统运行一段时间后才趋于稳定,在设计针对大型活动的监测系统时应该预留足够的试运行期。监测数据存在明显的周末与节假日效应,在选用短基线预警模型时需要充分考虑这一特点。

参 考 文 献

[1] Jorm L, Visotina M. The Sydney Olympics: a win for public health.

NSW Public Health Bull, 2003, 14(3):43-45.

- [2] Osaka K, Takahashi H, Ohyama T. Testing a symptom-based surveillance system at high-profile gatherings as a preparatory measure of bioterrorism. *Epidemiol Infect*, 2002, 129:429-434.
- [3] Tusi FC, Espino JU, Wagner MM, et al. Data, network and application; technical description of the Utah RODS Winter Olympic Biosurveillance System. *Proc AMIA Symp*, 2002; 815-819.
- [4] Zu RQ, Cai YS, Qin PZ, et al. The feasibility of outpatient data used for respiratory syndromic surveillance in general hospital. *Chin J Epidemiol*, 2010, 31(5):554-558. (in Chinese)
祖荣强, 蔡衍珊, 秦鹏哲, 等. 利用综合医院门诊病例数据开展呼吸道疾病症候群监测的探讨. *中华流行病学杂志*, 2010, 31(5): 554-558.
- [5] Marc AB, Kyle DM. ED syndromic surveillance for novel H1N1 spring 2009. *Am J Emerg Med*, 2011, 29:70-74.
- [6] Magruder SF, Lewis SH, Najmi A, et al. Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance. *MMWR*, 2004, 53(Suppl):S117-122.
- [7] Wei Z, Robert A, David JM, et al. Potential for early warning of viral influenza activity in the community by monitoring clinical diagnoses of influenza in hospital emergency departments. *BMC Public Health*, 2007, 7:250.
- [8] Dai XQ, Liu M, Kadeerbai HS, et al. Analysis of syndromic surveillance of 63 325 outpatients in fever clinic in Beijing. *J Peking University (Health Sciences)*, 2011, 43(3):375-378. (in Chinese)
代小秋, 刘民, 海山·卡德尔拜, 等. 北京市发热门诊63 325例就诊患者症状监测分析. *北京大学学报:医学版*, 2011, 43(3): 375-378.
- [9] CDC. Framework for evaluating public health surveillance systems for early detection of outbreaks. *MMWR*, 2004, 53(No. RR-5).

(收稿日期:2012-01-28)

(本文编辑:尹廉)