

采用北京市麻疹的暴发模拟数据比较 几种预警模型的功效

王小莉 王全意 刘东磊 曾大军 程贺 李素 段玮 黎新宇 栾荣生 贺雄

【摘要】 目的 优选出北京市麻疹的最佳预警模型及其参数,为其自动预警提供技术支持。方法 利用暴发模拟软件生成一系列不同性质的暴发信号,将其添加到北京市2005—2007年麻疹的实际日报数中。综合采用指数加权移动平均(EWMA)、C1-MILD(C1)、C2-MEDIUM(C2)、C3-ULTRA(C3)及时空重排扫描统计等预警模型识别加入的暴发信号,比较各模型不同参数的约登指数(YD指数)和检出时间(DT),优选出最佳模型参数,进而比较各模型最优参数下的预警功效,优选出最佳预警模型。结果 EWMA模型的最优参数为 $\lambda=0.6, k=1.0$; C1的最优参数为 $k=0.1, H=3\sigma$; C2的最优参数为 $k=0.1, H=3\sigma$; C3的最优参数为 $k=1.0, H=4\sigma$;时空重排扫描模型的最优参数为时间聚集性最大值为7 d,空间聚集性最大值为5 km。各模型的预警功效评价结果:EWMA的YD指数为90.8%, DT为0.121 d; C1的YD指数为88.7%, DT为0.142 d; C2的YD指数为92.9%, DT为0.121 d; C3的YD指数为87.9%, DT为0.058 d;时空重排扫描的YD指数为94.3%, DT为0.176 d。结论 5种模型中,时空重排扫描的预警功效最优。

【关键词】 麻疹; 预警; 指数加权移动平均; 时空重排扫描

Comparison between early outbreak detection models and simulated outbreaks of measles in Beijing WANG Xiao-li, WANG Quan-yi, LIU Dong-lei, ZENG Da-jun, CHENG He, LI Su, DUAN Wei, LI Xin-yu, LUAN Rong-sheng, HE Xiong. *Beijing Center for Disease Control and Prevention, Beijing 100013, China*

Corresponding author: HE Xiong, Email: hexiong@vip.163.com

【Abstract】 Objective Using simulated outbreaks to choose the optimal model and its related parameters on measles so as to provide technical support for developing an Auto Warning System(AWS). Methods AEGIS-Cluster Creation Tool was applied to simulate a range of unique outbreak signals. Then these simulations were added to the actual daily counts of measles from the National Disease Surveillance System, between 2005 and 2007. Exponential weighted moving average (EWMA), C1-MILD (C1), C2-MEDIUM (C2), C3-ULTRA (C3) and space-time permutation scan statistic model were comprehensively applied to detect these simulations. Tools for evaluation as Youden's index and detection time were calculated to optimize parameters before an optimal model was finally chosen. Results EWMA ($\lambda=0.6, k=1.0$), C1 ($k=0.1, H=3\sigma$), C2 ($k=0.1, H=3\sigma$), C3 ($k=1.0, H=4\sigma$) and space-time permutation scan statistic (maximum temporal cluster size=7 d, maximum spatial cluster size=5 km) appeared to be the optimal parameters among these models. Youden's index of EWMA was 90.8% and detection time being 0.121 d. Youden's index of C1 was 88.7% and detection time being 0.142 d. Youden's index of C2 was 92.9% and detection time being 0.121 d. Youden's index of C3 was 87.9% and detection time being 0.058 d. Youden's index of space-time permutation scan statistic was 94.3% and detection time being 0.176 d. Conclusion Among these five early warning detection models, space-time permutation scan statistic model had the highest efficacy.

【Key words】 Measles; Early outbreak detection; Exponential weighted moving average; Space-time permutation scan-statistic

2003年的SARS事件之后,使人们认识到建立

传染病预警系统的必要性和迫切性。国内外学者开展了一系列传染病预警模型研究^[1-7],但均未在模型预警功效评价前优选模型参数。本研究采用AEGIS-CCT软件模拟暴发信号^[8],在优选出模型参数的基础上进一步比较各模型最优参数下的预警功效,使预警功效的评价更为客观准确。

DOI: 10.3760/cma.j.issn.0254-6450.2009.02.015

基金项目:北京市自然科学基金资助项目(7082047);北京市优秀人才培养资助项目(2007id0302600112)

作者单位:100013北京市疾病预防控制中心(王小莉、王全意、刘东磊、段玮、黎新宇、贺雄);中国科学院自动化研究所(曾大军、程贺、李素);四川大学华西医学中心公共卫生学院(栾荣生)

通信作者:贺雄,Email: hexiong@vip.163.com

资料与方法

1. 数据来源及分类: 根据实际情况, 本研究以街道为基本地理单位。从“疾病监测信息报告管理系统”中获取北京市 2005—2007 年各街道麻疹的实际日报数。其中 2005—2006 年的实际数据作为筛选数据库的基线数据, 而将 2007 年的实际数据作为优选数据库的基线数据。

2. 方法和内容:

(1) 数据的拟合: 以 2005 年底卫生部应急办公室出台的《国家突发公共卫生事件相关信息报告管理工作规范(试行)》中麻疹的报告标准作为暴发标准, 剔除实际存在的暴发信号, 继而利用 AEGIS-CCT 软件按既定标准设置暴发信号参数。将暴发病例数设为 10 例, 持续时间设为 1~7 d, 病例的时间分布模型分别为随机型、直线型和指数型。利用 JAVA 编程, 从 300 个街道(乡镇)中随机抽取一定数量的街道, 并在分析的整个时间段内随机选取 1 d 作为暴发的起始日期, 从暴发起始日期开始, 在选取的街道中逐一将生成的暴发信号添加到筛选数据库的基线数据中合成筛选数据库, 用于筛选模型的备选参数; 而将模拟信号添加到优选数据库的基线数据中合成优选数据库, 用于从备选参数中优选出最佳预警模型参数。为控制抽样误差, 重复抽样 10 次, 以均值作为结果。

(2) 评价指标及优选方法: 评价指标包括 YD 指数(Youden's index)和检出时间(detection time, DT)^[9]。YD 指数, 即灵敏度与特异度之和减去 1, 指数范围在 -1~1 之间。指数越大, 模型识别暴发的能力越强。YD 指数综合考虑了灵敏度和特异度的总效应, 反映预警模型识别真实暴发与非暴发的总能力。DT, 即模型能在暴发发生后将其识别出来的时间, 反映了预警模型识别暴发的及时性。

优选方法大体上以 YD 指数最高者作为最佳结果, 同时参考各自在 DT 上的差异大小。当 YD 指数相等或相差不大时(差值 < 5.0%), 而 DT 之间的差值 > 0.5 d, 可优先考虑 DT 较短者。比较不同参数组合下的 YD 指数和 DT, 按优选原则选出最佳参数; 进而比较各模型最优参数下的 YD 指数和 DT, 按优选原则, 选出最佳预警模型。

(3) 指数加权移动平均(exponentially weighted moving average, EWMA): 设过程观测值 $X_t \sim N(\mu, \sigma^2)$, 则 EWMA 的表达式为:

$$Z_t = \lambda X_t + (1 - \lambda) Z_{t-1} \quad (1)$$

$$UCL = \mu + k \frac{\hat{\sigma}}{\sqrt{n}} \sqrt{\left(\frac{\lambda}{2 - \lambda}\right) [1 - (1 - \lambda)^{2n}]} \quad (2)$$

当 $Z_t > UCL$ 时, 则认为发病水平存在异常。模型中两个参数为权重因子 $\lambda (0 < \lambda < 1)$ 和控制限参数 k ^[10, 11]。根据文献[12, 13]中 k 的取值范围, 将其限定为 $0 < k \leq 3$ 。参数调试方式为: 每间隔 0.1 调节 λ , 每间隔 0.5 调整 k 。采用移动标准差 s 作为 σ 的估计值, 移动均值 MA 作为 μ 的估计值。移动平均周期以研究病种的潜伏期长短为参考, 将麻疹的移动平均周期设定为 14 d。

(4) 累计和(cumulative sum): 一般认为研究的特征参数 $X \sim N(\mu, \sigma)$, 令:

$$C_t = \max\{0, X_t - (\mu + k\sigma) + C_{t-1}\} \quad (3)$$

式中初始值 $C_0 = 0$; $k\sigma$ 为允偏差; 若均值由 μ 偏移到 $\mu + k\sigma$ 时, 希望能引起警觉; H 为判定值, 当 $C_t \geq H$, 认为增加有统计意义^[14]。根据其识别异常的灵敏度高, 将 C_t 分为三类, 分别为 C1-MILD (C1)、C2-MEDIUM (C2) 和 C3-ULTRA (C3)。移动平均周期为 14 d, C1、C2、C3 的计算公式如下。

$$C1 = \max\{0, X_t - (MA_1 + kS_1) + C_{t-1}\} \quad (4)$$

$$C2 = \max\{0, X_t - (MA_2 + kS_2) + C_{t-1}\} \quad (5)$$

C3 等于 C2 计算公式中求得的 C_t 、 C_{t-1} 和 C_{t-2} 之和。其中 MA_1 和 S_1 分别代表 $t-14$ 至 $t-1$ 期间报告病例数的移动均值和移动标准差, MA_2 和 S_2 分别代表 $t-16$ 至 $t-3$ 期间报告病例数的移动均值和移动标准差, 忽略了最近 2 d 的数据。

根据既往文献中 H 和 k 的取值范围^[15-17], 将其限定为 $3\sigma \leq H \leq 5\sigma, 0 < k \leq 1.5$ 。参数调试方式为: 每间隔 0.1 调节 k 的大小, 每间隔 0.5 σ 调整 H 。采用移动标准差 S 作为 σ 的估计值, 移动均值 MA 来作为 μ 的估计值。

(5) 时空重排扫描统计: 假定将整个研究区域分为若干个子区域 Z , 整个研究时期为 $T, 0 < d < T$, 子区域 Z 在 d 天中的发病数用 $C_{z,t}$ 表示, 整个区域在总时间 T 内的总发病数为 C ; 每个子区域内的日期期望发病数为 $\mu_{z,t}$, 每个圆柱 A 中对应的期望发病数为 μ_A 。

令 C_A 为扫描圆柱 A 中的实际发病数, 其概率函数为:

$$P(C_A) = \frac{\left(\sum_{z,t} C_{z,t}\right) \left(C - \sum_{z,t} C_{z,t}\right)}{\left(\sum_{z,t} C_{z,t}\right)^2} \quad (6)$$

采用广义似然函数计算实际发病数和理论发病数的比值 LLR, 选出 LLR 最大的窗口。利用蒙特卡罗法产生模拟数据集, 计算 P 值, 判断该窗口的聚集度是否具有统计学意义。

$$LLR = \frac{L(A)}{L_0} = \left(\frac{C_A}{\mu_A}\right)^{c_A} \left(\frac{C - C_A}{C - \mu_A}\right)^{c - c_A} \quad (7)$$

其扫描窗口为圆柱形, 圆柱的底对应一定的地理区域, 圆柱的高对应一定的时间长度。扫描窗口的大小和位置处于动态的改变之中^[18]。最长聚集时间的确定常以研究病种的潜伏期为参考, 麻疹的平均潜伏期为 9~12 d, 因此本研究将最长聚集时间参数设定为 (3 d, 7 d, 10 d, 12 d)。最大聚集空间半径可按辖区区域大小划分, 也可按人口的比例划分, 由于缺乏各街道的人口数, 因此空间半径按辖区区域大小划分, 结合各辖区的区域半径, 最终设定的空间半径为 (2 km, 5 km, 8 km, 10 km)。

(6) 统计学分析: 本研究拟综合采用 SPSS 13.0、EXCEL、JAVA 编程、AEGIS-CCT 及 SaTscan 软件对数据进行分析 and 利用。采用 SPSS 13.0 软件进行数据整理和描述统计, 利用 JAVA 编程进行指数加权移动平均和累计和移动平均分析, 判别发病水平是否存在异常。利用 SaTscan 软件进行时空重排扫描统计分析, 分析北京市各街道病例聚集性, 判断其发病水平是否存在异常。

结 果

1. 参数优选: EWMA 模型优选出的参数组合为 $\lambda = 0.6, k = 1.0$; C1 模型优选出的参数组合为 $k = 0.1, H = 3\sigma$; C2 模型优选出的参数组合为 $k = 0.1, H = 3\sigma$; C3 模型优选出的参数组合为 $k = 1.0, H = 4\sigma$; 时空重排扫描模型优选出的参数组合为时间聚集性的最大值为 7 d, 空间聚集性的最大值为 5 km。

2. 预警功效评价: 比较各模型最优参数下的预警功效, 优选出最佳预警模型, 各模型的预警功效见表 1。6 种预警模型中, 时空重排扫描模型的平均 YD 指数最高, 其次为 C2, 而 C3 的 YD 指数最低; C3 的 DT 最短, 而时空重排扫描的 DT 相对较长。6 种模型中, 根据既定的优选原则, 时空重排扫描模型与 C3 模型 YD 指数之间差值 $> 5.0\%$, 而 DT 之间的差值 $< 0.5 d$ 。相对于时空重排扫描 94.3% 的 YD 指数, 可忽略 C3 模型在 DT 方面的优势。最终认为对于麻疹来说, 时空重排扫描模型的预警功效最高。

表 1 麻疹不同模型的预警功效

预警模型	YD 指数 (%)			DT (d)		
	最大值	均值	最小值	最大值	均值	最小值
EWMA(0.6, 1.0)	92.2	90.8	79.1	0.278	0.121	0.037
C1 (0.1, 3σ)	91.9	88.7	85.5	0.364	0.142	0.076
C2(0.1, 3σ)	96.0	92.9	90.3	0.228	0.121	0.048
C3 (1.0, 4σ)	89.7	87.9	75.2	0.124	0.058	0.032
时空重排扫描(7 d, 5 km)	98.9	94.3	91.3	0.325	0.176	0.180

讨 论

由于大多数传染病暴发频率较低, 甚至较长时间内无暴发, 使预警模型的功效评价难以进行。为解决这一问题, Buckridge 等^[19, 20]和 Hutwagner 等^[21, 22]先后采用模拟暴发信号的方法对一些预警模型的功效进行评价, 开创了预警模型研究的新思路。本研究应用软件模拟暴发信号, 在优选出最佳模型参数的基础上评价模型预警功效, 其评价结果更为客观有效。

不同的模型参数, 其预警效果亦不同。研究结果证实, 同一模型, 采用不同的参数, 其预警功效差异较大。EWMA (0.6, 1.0) 的 YD 指数为 88.9%, DT 为 0.086 d。而 C1 (1.5, 5σ) 的 YD 指数为 27.1%, DT 为 0.324 d, 若比较 EWMA 与 C1 的预警功效, EWMA (0.6, 1.0) 在 YD 指数和 DT 上均优于 C1 (1.5, 5σ) 的预警功效, 从而得出 EWMA 模型的预警功效优于 C1 模型的结论。而实际上, C1 (0.1, 3σ) 的灵敏度为 92.9%, 特异度为 97.5%, YD 指数为 90.4%, DT 为 0.028 d, YD 指数和 DT 均优于 EWMA (0.6, 1.0)。两者的预警功效差异在很大程度上是由参数选取不当造成的, 并非模型内在的差异引起。因此, 评价模型的预警功效前, 需优化模型参数, 优选出模型的最佳参数, 进而比较各模型最佳参数下的预警功效, 才能客观选出最优模型。

本研究存在局限性。由于病例数据来源于以病例为基础的“疾病监测信息报告管理系统”。目前网络直报的工作环节中, 仍然存在临床诊断标准和报告规范不统一, 现住址和工作单位不准确或不完整, 部分病例的报告不及时等问题, 一定程度上影响了直报数据的准确性、完整性和及时性, 势必会影响预警分析的准确性及有效性。良好的数据质量是统计分析的基础, 因此有必要加强对网络直报数据的质量管理, 统一诊断标准和报告规范, 完善相关地址 (居住地与工作或学习单位) 填报制度, 提高直报人员的工作意识。

研究给出各种模型的最优参数组合及模型各参数及评价指标值的变化规律与文献报道基本吻合^[9, 20, 22, 23]。各模型的灵敏度、特异度和检出时间略高于既往研究的平均水平,究其原因可能为:与以往研究不同,本研究计算出的预警功效为各模型最优参数下的水平,预警功效相对较高。此外,还可能与既往研究多以症状监测数据为研究对象有关。症状监测数据与以疾病为基础的病例数据相比,其特异性不高,准确性较低,可能导致灵敏度和特异度降低。

本次研究结果表明,对于麻疹,时空重排扫描统计模型的预警功效最高。时空重排扫描综合考虑了时间和空间两个因素,其扫描窗口的大小和位置不断变化,避免了选择偏向;同时还配有专门的统计分析软件,参数调整及运算相对简单。因此,认为时空重排扫描统计模型在麻疹自动预警方面具有广阔的应用前景,有待于进一步开发和运用。

参 考 文 献

- [1] Hutwagner L, Thompson W, Seeman GM, et al. The bioterrorism preparedness and response Early Aberration Reporting System (EARS). *J Urban Health*, 2003, 80(2 Suppl 1):S89-96.
- [2] Yih WK, Abrams A, Danila R, et al. Ambulatory-care diagnoses as potential indicators of outbreaks of gastrointestinal illness-minnesota. *MMWR*, 2005, 54(Suppl):S157-162.
- [3] Yih WK, Caldwell B, Harmon R. The national bioterrorism syndromic surveillance demonstration program. *MMWR*, 2004, 53 (Suppl): S43-46.
- [4] Wong WK, Moore A, Cooper G, et al. WSARE: What's strange about recent events? *J Urban Health*, 2003, 80 (2 Suppl 1): S66-75.
- [5] 杨维中,邢慧娟,王汉章,等.七种传染病控制图法预警技术研究. *中华流行病学杂志*, 2004, 25(12): 1039-1041.
- [6] 马家奇,王丽萍,戚晓鹏,等.基于网络直报的传染病监测自动预警信息系统概念模型. *疾病监测*, 2006, 21(12): 679-681.
- [7] 殷菲,冯子键,李晓松,等.基于前瞻性时空重排扫描统计量的传染病早期预警系统. *卫生研究*, 2007, 36(4): 455-458.
- [8] Christopher AC, Karin I, Karen LO, et al. A software tool for creating simulated outbreaks to benchmark surveillance systems. *BMC Med Inform Decis Mak*, 2005, 5:22.
- [9] Hutwagner L, Browne T, Seeman GM, et al. Comparing aberration detection methods with simulated data. *Emerg Infect Dis*, 2005, 11:314-316.
- [10] Hunter JS. The exponentially weighted moving average. *J Qual Technol*, 1986, 18(4):203-210.
- [11] Michael LJ, Atar B, Ian P, et al. A simulation study comparing aberration detection algorithms for syndromic surveillance. *BMC Med Inform Decis Mak*, 2007, 7:6.
- [12] 王海宇,徐济超,杨剑锋,等.基于 APL 的 EWMA 图优化设计. *系统工程理论方法应用*, 2006, 15(1):65-70.
- [13] 吉明明,孙浩.具有可变抽样区间的 EWMA 标准差控制图. *数学的实践与认识*, 2007, 37(12):90-96.
- [14] Montgomery DC. *Introduction to statistical quality control*. 3rd ed. John Wiley and Sons Inc, 1996: 314-330.
- [15] Carpenter TE. Evaluation and extension of the cusum technique with an application to Salmonella surveillance. *J Vet Diagn Invest*, 2002, 14:211-218.
- [16] 何桢,刘豹.累计和控制图在测量系统稳定性控制中的应用研究. *系统工程学报*, 2002, 17(6):547-550.
- [17] 王殊秩. CUSUM 控制图的一种优化设计方法研究. *工业工程与管理*, 2004, 2:44-46.
- [18] Kulldorff M, Heffernan R, Hartman J, et al. A space-time permutation Scan statistic for disease outbreak detection. *PLoS Medicine*, 2005, 2(3):216-224.
- [19] Buckeridge DL, Burkom H, Moore A, et al. Evaluation of syndromic surveillance systems-design of an epidemic simulation model. *MMWR*, 2004, 53(Suppl):S137-143.
- [20] Buckeridge DL, Switzer P, Owens D, et al. An evaluation model for syndromic surveillance: assessing the performance of a temporal algorithm. *MMWR*, 2005, 54(Suppl):S109-115.
- [21] Hutwagner LC, Thompson WW, Groseclose S, et al. An evaluation of alternative methods for detecting aberrations in public health surveillance data. *American Statistical Association Joint Statistical Meetings, Proceedings of the Biometrics Section: Indianapolis*, 2000:82-85.
- [22] Hutwagner LC, Thompson WW, Seeman GM, et al. A simulation model for assessing aberration detection methods used in public health surveillance for systems with limited baselines. *Stat Med*, 2005, 24:543-550.
- [23] Mandl KD, Reis B, Cassa C. Measuring outbreak-detection performance by using controlled feature set simulations. *MMWR Morb Mortal Wkly Rep*, 2004, 53 Suppl:S130-136.

(收稿日期:2008-07-31)

(本文编辑:张林东)