

# 控制图原理及在传染病预警中的应用现状

秦利平, 马家奇

**摘要:** 在传染病发生的早期阶段做出预警,可及时采取干预措施,尽可能降低传染病对人类健康和社会经济、政治等的影响。控制图法是一种简单易行且功效高的预警技术,但也有不足之处。

**关键词:** 控制图;传染病;预警

中图分类号: R51

文献标识码: A

文章编号: 1003-9961(2008)02-0117-04

**Use of control chart in early warning of infectious disease** QIN Li-ping, MA Jia-qi. Center for Public Health Surveillance and Information Service, Chinese Center for Disease Control and Prevention, Beijing 100050, China

**Corresponding author:** MA Jia-qi, Email: majq@chinacdc.cn

**Abstract:** Warning signals in the early phase of communicable diseases could facilitate timely intervention, so as to do less harm to the public health, as well as social economics and politics. Control chart has been regarded as a practical and efficient way for disease warning, with much more to be improved.

**Key words:** control chart; infectious disease; early warning

传染病预测是根据传染病的发生、发展规律及有关因素,用分析判断和数学模型等方法对可能发生的传染病的流行趋势做出分析和预判,是制定预防和控制传染病应对策略的前提<sup>[1]</sup>。传染病预测的方法有定性预测、定量预测和综合预测。定性预测中有流行控制图法、比数图法、专家预测法、德尔菲法(Delphi),逐步判别分析、模糊聚类预测法、贝叶斯概率判断法及尤度法。定量预测中有回归预测法、灰色预测模型、时间序列平滑法、趋势外推法、马尔科夫链预测方法、博克斯-詹金斯模型(B-J 模型)<sup>[1,2]</sup>。

信息量的不断丰富提升了在更早的阶段识别当地疾病暴发的希望,以至于暴发一开始就能被发现和包含在各种信息中,从而为实现预警提供了可能。预警是利用预测技术,在考虑了资料的不完全性和危害的不确定性之后,及早发现传染病异常变化的苗头,发出警示,提醒工作人员及时调查核实,以达到早发现早处理的目的<sup>[1]</sup>。传染病预警是预测技术在实际中的重要应用。用于预警的预测技术,应当能体现短期化、简单化、经常化和模型自动化的特点<sup>[1]</sup>。

预警方法是传染病预警研究的重要内容之一,对预警系统的设计、改进及预警效果影响因素的分析均具有重要的意义。本文在概述预警方法的基础上,对控制图原理及其在传染病预警中的应用情况进行分析和综述。

## 1 预警方法的概述

根据不同的资料来源和数据类型,可使用不同的预警方法。疾病的定性预警可采用综合预测法、控制图法、尤度法、Bayes 概率法、逐步判别分析等多种统计方法,实现对疾病的发展趋势和强度的定性估计,明确是上升还是下降,是流行还是散发。疾病的定量预警可采用直线预测模型、指数曲线预测模型、多元逐步回归分析、GM 灰色模型预测、简易时间序列季节周期回归模型等方法<sup>[3]</sup>。

有研究表明,基于自动化的实验室监测系统,使用分级时间序列模型可早期自动识别传染病的暴发。当实际发生值与预期值比较超过预警界值水平时,系统将发出警报,用模拟序列和观察序列均证明系统有较好的灵敏度、特异度,且及时性好<sup>[4]</sup>;而基于时空扫描的统计学方法识别较大范围内暴发的能力较高但发现某个局限范围内暴发的能力较低<sup>[5]</sup>。

除了现有疾病监测系统,还有其他的多种数据来源,如药房销售记录、学校(工厂)因病缺课(勤)记录、居民医疗保险等记录,这些多来源的数据可使用

基金项目:1.国家自然科学基金(编号:30590371);2.国务院应急管理办公室“十一五”支撑课题

作者单位:中国疾病预防控制中心,北京 100050

作者简介:秦利平,女,四川省人,中国疾病预防控制中心在读 MPH 硕士,目前参与中国疾病预防控制中心公共卫生监测与信息服务系统传染病预警信息系统课题研究

通讯作者:马家奇, Tel:010-63174191, Email: majq@chinacdc.cn

收稿日期:2007-04-18

高斯混合模型和 Bayes 概率模型去早期识别异常的发生,达到早期预警和优化资源配置的目的<sup>[6]</sup>。

## 2 控制图原理

**2.1 控制图的定义** 控制图是由美国 Walter A. Shewhart 于 1924 年首创。控制图(control chart)也叫统计过程控制(statistical process control, SPC),是用于辨别普通原因和特殊原因导致的过程变化的一种统计工具<sup>[7]</sup>,利用历史数据来确定控制的中心线(平均线)、上控制限和下控制限。将目前的数据与这些限值做比较,可以得出过程变化是否在控制范围内的结论。控制图也是统计质量控制(Statistical quality control, SQC)的工具之一,在技术上是 SQC 最精确的工具<sup>[8]</sup>。

**2.2 控制图的类型** 常用的控制图有三类:一是 Shewhart 控制图,最常用,但在过程均值有小的漂移时,Shewhart 控制图的效果不是很好;二是累积和控制图(CUSUM),基于似然比导出;三是指数加权移动平均控制图(EWMA),后两种已被证明在检测小的漂移时效果不错。CUSUM 在与 EWMA 的比较中更具优势,应用更广<sup>[9]</sup>。本文仅讨论 Shewhart 控制图。

不同的数据使用不同类型的控制图,而不同类型的控制图建立在不同的概率模型的基础上<sup>[10]</sup>。控制图的类型与分类方法有关,按变量多少可分为单变量控制图和多变量控制图;按资料的类型,常用的控制图有:计数资料用 XmR(individual X and moving range chart)控制图;计量资料的控制图类型与样本量有关,可用 XmR 控制图和  $\bar{X}$ -Bar and R 控制图(样本量大时)<sup>[11]</sup>。根据控制对象的不同,可分为变量控制图(如: $\bar{X}$  and R chart; $\bar{X}$  and S chart;chart of individuals;moving average-moving range chart, etc)和属性控制图(如:p 图、np 图、u 图和 c 图)<sup>[7,12]</sup>。对任何一种资料类型,控制图可分为短期运行图(short term charts,又叫 Z 图)和组图(group charts),后者又叫多重属性图(multiple characteristic charts)<sup>[7]</sup>,有学者认为,一个多重属性控制图通过提供核对表(或运算表)成为最简单的传统问题的解决工具。当过程失去控制时,附加的表格程序可提供一个警报<sup>[13]</sup>,为自动预警的实现提供了可能。

有研究认为,基于中位数分析的 Tukey's 控制图,不受假设数据分布的限制,能应用于小数据集,不受极端值的影响,不需要计算标准差或平均数<sup>[14]</sup>,从而使这种控制图的应用更加方便。

SPC 图还有一种类型,被称为 g 和 h 控制图,用于监控一些不良事件,如院内感染、心脏手术后的并发症等,效果优于传统的方法,尤其是识别罕见事件过程中率的变化<sup>[10]</sup>。

**2.3 控制图的特性** 所有的 SPC 控制图有以下的特性<sup>[15]</sup>:a. X 轴是连续的,通常是表示时间进程;b. Y 轴是每个点在时间上以图表的形式表示的统计数值;c. 控制限通过统计学定义并被描绘在图上,这些统计学的控制限由以前的和/或目前的过程数据来决定;d. 过程中观察到的变化,如为发生在控制限之间点的波动,认为是过程内在的变化,可以把这种变化归结于“普通原因”;e. 当一些点超过了控制限,则认为过程已经发生了变化。控制限之外的任何点可被归因于某个“特殊原因”。

## 3 控制图应用于预警时应关注的问题

**3.1 控制限(control limits)的设定** 控制限的设定影响到了异常信号的判定。控制限主要有 3 种设定方式:指定控制限、按照理论值计算、按照公式计算。控制图中有 3 条曲线,分别代表上警戒线、下警戒线和中位数线,而警戒线是可以调整的<sup>[11]</sup>。当以明显异常作为警戒目标时,警戒线即为控制限。如果仅对增加的发生水平感兴趣,在控制图方法的应用中可以不使用下控制限<sup>[16]</sup>。

上下控制限的计算与资料的分布类型有关,用于计算的公式是常规概率理论的直接应用。用中心线到加/减 3 个标准差进行控制限的设置,仅适合于正态分布或分布形式与正态分布相似时<sup>[15]</sup>。对于其他类型的分布,可能会对这些信号做出更频繁的反应,即使过程没有发生变化或变化较少时<sup>[15]</sup>,最终可能出现较多的假预警。控制限的计算还可以依靠二项式标准的近似法和泊松分布<sup>[13]</sup>。

中心线(center line)的计算也与资料的分布类型有关。当资料不符合正态分布时,中心线即为中位数线。对于控制图而言,中心线(即基线水平)通过一个原始的数据集来决定,流行期间的数据应该被排除在计算之外。参考的基线越稳定,设定的控制限越精确。在这样的基线测量方法不可得的情况下,可以使用历史的数据<sup>[16]</sup>。

**3.2 窗口期的设定** 识别系统使用的窗口期(temporal window)对不同强度的暴发识别可能有重要的影响。目前研究中,窗口期的设定并不统一。有研究用观察月既往 5 年该月及其前后一个月<sup>[17]</sup>;有研究用疾病的平均潜伏期并做前后摆动<sup>[18]</sup>;还有研究用 7 天,

但是识别的窗口期定为 7 天的移动平均数,在某个特定的条件下将会平滑持续 2~3 d 的暴发,这种平滑可能降低暴发的信号。相反地,用一个较短的移动窗口可能不会发现在数量上缓慢增加的暴发<sup>[19]</sup>。

3.3 数据来源及其形式 设定控制限时,有研究用月发病率<sup>[17]</sup>。国外一项研究证明,用周病例数作控制图的这种简单的算法与其他复杂的算法相比,在发现疾病流行的能力上效果一样的好,而且可用于检验其他的预警界值算法。用周病例数的某个百分位数作预警界值,由于这种方法对极端值不敏感,故其适用范围扩大<sup>[20]</sup>。一项监控重大活动期间传染病等事件的研究,由于监控持续的时间有限,故使用了日报告病例数<sup>[16]</sup>。虽然用媒介、气象等监测结果作寄生虫病的预警可能会比用周病例数作预警的时间提前,但这种基于气象的监测系统在分析时不得不依赖相对复杂的算法,从而使这种预警方法难以实施和开展<sup>[20]</sup>。

4 控制图在传染病预警中的应用

无论测量的是什么过程,均可用 SPC 来分析过程中的变化<sup>[15]</sup>。SPC 已经广泛地应用于医学、环境监测、制药、卫生、食品、教育等领域。有学者将控制图用于传染病监测及报卡的质量管理,以便判断是否存在传染病暴发或流行、判断辖区内医疗单位是否有传染病的漏报或集中报告,见图 1。

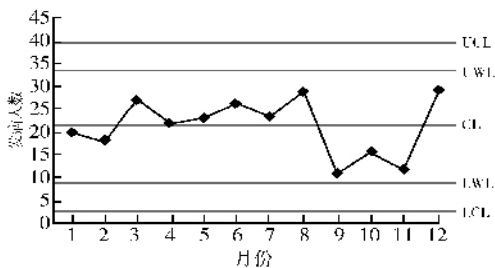


图 1 乙肝疫情质量控制图  
Figure 1 Quality control of type B hepatitis

传染病预测预警方面,流行控制图法适用于各种分布的传染病,对于有季节性流行或周期性流行规律的传染病,效果较好,是一种较好的预警方法。

美国从 1990 年 4 月开始,使用图解的方式来表示法定传染病的监测数据,与基线数据作比较,以发现可能的暴发,并使用这种方法成功地识别了发生在加州的成人风疹的暴发<sup>[21]</sup>。另一项研究以均数加两个标准差为控制限,回顾性地验证了控制图在识别登革热暴发中的作用<sup>[22]</sup>。

对世界杯等人群大规模聚集活动进行动态监控

时,国外一项研究也成功地应用了控制图。通过设立监测哨点,由临床医生每日报告收集数据,使用计算均数和限值的方法、与参考的基线时期做比较,决定被评估的过程是否稳定,最终实现传染病预警的目的<sup>[16]</sup>。该研究认为,控制图提供了可被内科医生、公共卫生工作者甚至是非专业的公众轻松解释的信息,并有可能鉴别疾病固有的随机变化和非随机变化,而后者是一个流行警报的潜在信号。

国内有学者利用传染病(菌痢、伤寒)历年月发病率最大值、最小值和中位数分别划出上控制限、下控制限和中位数线,根据该流行曲线来推测传染病发生或流行的趋势或强度<sup>[23]</sup>。当同期发病率呈正态分布时,用均数±3 倍标准差作为上、下控制限;当同期发病率呈对数正态分布时,可用几何均数和几何标准差来作图。还有学者利用传染病(流脑)历年月发病率资料,以标化月均发病率和标化月最高发病率作控制图,预测传染病的发病趋势。

国内一项研究以 2004 年各病种(发病数>20 例、发病数符合或近似正态分布)的周报告发病数(从中剔除周暴发病例数)为基线数据,计算 2 周移动平均数和移动平均数标准差,再利用公式计算预警值;用基线数据、预警值、2005 年数据作为控制图的 3 个数据系列作图,成功地用于传染病监测预警。

国内另一项研究按照控制图法预警模型原理,以研究疾病观察月既往 5 年该月及其前后一个月的月发病率数据为基线数据,用专家咨询法确定流行参照标准,优选出 7 种传染病的合适预警界值<sup>[17]</sup>,同时证明用 5 年基线数据建模,将预警时间单位向前和向后各移动一个单位作为基线数据的处理方式是可行的。

还有一项研究采用移动百分位数的思路,在动态预警控制图方法的基础上,对基于网络直报的传染病监测实现以日为单位的动态自动预警提出了信息系统实现的概念模型,并通过监测信息平台配置调节参数,动态校正预警阈值的灵敏度和特异度,以达到探测传染病暴发早期信息的目的<sup>[18]</sup>,见图 2。

5 结论

理论上,用于预警的方法较多,但大都依赖于相对复杂的统计学模型。与其他预警算法相比,控制图法是一种简单易行且功效高的预警技术,但其在应用中存在一定的局限,主要表现在:a 利用历史数据作为基线资料,而自然暴发疾病的数据被掩盖在历史数据之中;b 不适宜的控制限的设定会使假阳性



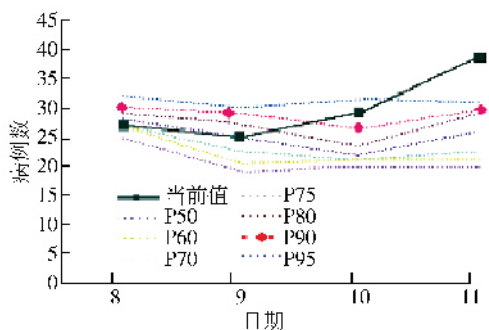


图 2 某传染病的动态百分位数控制图

Figure 2 Control chart for dynamic percentile of a certain infectious disease

增加。研究认为控制限越窄,产生错误预警的可能性越大,如使用均数加减 2 个标准差作为控制限时,比使用均数加减 3 个标准差作为控制限产生的假预警的机会更多;c 用敏感度、特异度、阳性预测值等指标评价其预警效果时,如何确定金标准,这也是评价其它预警算法时共存的问题;d 控制图法更多地关注数据的时间属性特征,而对数据的空间属性特征考虑不够,无法同时实现时间和空间两个维度的预警。

目前,无论是卫生行政管理者还是传染病相关专业机构,都对传染病预警十分关注。国内将控制图法应用于传染病预警的实例并不多,如何利用不同来源的数据(如基于病例报告、实验室检测或症状监测等来源)、如何选择适宜的控制限值和预警阈值、如何分析预警效果的影响因素等都有待于进一步探讨。

#### 参考文献

- [1] Xing HX, Yang WZ, Wang HZ. Forecast of infectious disease[J]. Journal of Preventive Medicine Information, 2004, 20(6): 639–642. (in Chinese)  
邢慧娟, 杨维中, 王汉章. 传染病预测[J]. 预防医学情报杂志, 2004, 20(6): 639–642.
- [2] Wang QD, Wang ZJ, Liu RF. Statistics forecast in modern health administration[J]. Chinese Journal of Health Statistics, 2001, 18(4): 245–248. (in Chinese)  
王启栋, 王贞洁, 刘荣甫. 现代卫生事业管理中的统计预测[J]. 中国卫生统计, 2001, 18(4): 245–248.
- [3] Wu ZM, Su L, Duan DH. A study on the model of the disease Control forecasting system [J]. China Public Health, 2003, 19(10): 1154–1155. (in Chinese)  
吴泽明, 宿鲁, 段德辉. 疾病控制预警系统模式探讨[J]. 中国公共卫生, 2003, 19(10): 1154–1155.
- [4] Heisterkamp SH, Dekkers AL, Heijne JC. Automated detection of infectious disease outbreaks; hierarchical time series models[J]. Stat in Medicine, 2006, 25(12): 4179–4196.
- [5] Martin K, Zhang Z, Hartman J, et al. Benchmark data and power calculations for evaluating disease outbreak detection methods [J]. MMWR, 2004, 53(9): 144–151.

- [6] Thomas DM, Arouh S, Carley K, et al. Automated anomaly detection processor for biologic terrorism early detection –hampton [J]. MMWR, 2005, 54(8): 203–203.
- [7] Data Collection And Analysis Tools. <http://www.asq.org/learn-about-quality/data-collection-analysis-tools/overview/control-chart.html>
- [8] Control Charts as a tool in SQC. <http://deming.eng.clemson.edu/pub/tutorials/qctools/ccmain1.htm>.
- [9] Pu XL. The improve of cumulative sum testing[J]. Journal of applied mathematics, 2003(2): 225–241. (in Chinese)  
濮晓龙. 关于累积和(CUSUM)检验的改进[J]. 应用数学学报, 2003(2): 225–241.
- [10] Benneyan JC. Number-between g-type statistical quality control charts for monitoring adverse events [J]. Health Care Mannt Sci, 2001, 4(4): 305–318.
- [11] Basic tools for process improvement, module 10: control chart. <http://www.balancedscorecard.org/files/opdef.pdf>.
- [12] Quality Control Charts. <http://www.statsoft.com/textbook/stquacon.html>.
- [13] Levinson WA. Quality control charts control multiple attributes [J]. ProQuest Health Man, 2004, 43(9): 40–43.
- [14] Aleml F. Tukey's Control Chart[J]. Q Manage Health Care, 2004, 13(4): 216–221.
- [15] SPC Concepts. [http://www.qualityamerica.com/knowledgecente/knowctrSPC\\_Concepts.htm](http://www.qualityamerica.com/knowledgecente/knowctrSPC_Concepts.htm).
- [16] Hanslik T, Boelle PY, Flahault A. The control chart: an epidemiological tool for public health monitoring[J]. Public Health, 2001, 115: 277–281.
- [17] Yang WZ. A study on early detection for seven infectious diseases [J]. Chinese Journal of Epidemiology, 2004, 12(25): 12–14. (in Chinese)  
杨维中. 七种传染病控制图法预警技术研究[J]. 中华流行病学杂志, 2004, 12(25): 12–14.
- [18] Ma JQ, Wang LP, Qi XP, et al. The conceptual model of the automatic early-warning system of infectious diseases monitoring based on direct network reporting system [J]. Dis Sur, 2006, 21(12): 679–681. (in Chinese)  
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[J]. 疾病监测, 2006, 21(12): 679–681.
- [19] Mandl KD, Reis B, Cassa C. Measuring outbreak-detection performance by using controlled feature set simulations [J]. MMWR, 2004, 53(9): 130–136.
- [20] Hailay DT, Joel S, Awash T, et al. Alert threshold algorithms and malaria epidemic detection, <http://www.cdc.gov/ncidod/eid/vol10no7/03-0722.htm>.
- [21] Current Trends Update: Graphic method for presentation of notifiable disease data -- United States, 1991, 40(7): 124–25. <http://www.cdc.gov/mmwr/preview/mmwrhtml/00001917.htm>.
- [22] Rigau-Perez JG, Millard PS, Walker DR, et al. A deviation bar chart for detecting dengue outbreaks in Puerto Rico [J]. Am J Public Health, 1999, 89(3): 374–378.
- [23] Zhang ZQ, Zhang YX. Prevalence control graph method [J]. Dis Surv, 1994, 9(1): 17–18. (in Chinese)  
张珍麒, 章扬熙. 流行控制图法[J]. 疾病监测, 1994, 9(1): 17–18.