

· 疫情监测 ·

七种传染病控制图法预警技术研究

杨维中 邢慧娟 王汉章 兰亚佳 孙乔 胡世雄 吕伟 袁政安 陈裕旭 董柏青

【摘要】 目的 为传染病预警系统提供预警方法和预警界值及提高传染病疫情监测的预警能力。**方法** 使用历史传染病报告数据建立预警数据库,采用控制图法建立预警模型,专家咨询法确定流行参照标准,通过计算、比较灵敏度、特异度、阳性预测值和绘制 ROC 曲线,选出合适的预警界值。**结果** 预警肾综合征出血热(出血热)、甲型肝炎(甲肝)、细菌性痢疾(痢疾)、流行性脑脊髓膜炎(流脑)、疟疾,选用 P_{80} 作为预警界值预警功效较好,灵敏度和特异度均在 90% 以上;预警麻疹、肺结核,选用 P_{90} 作为预警界值预警功效较好,灵敏度和特异度均在 85% 以上。**结论** 控制图法可以作为该 7 种传染病预警方法,出血热、甲肝、痢疾、流脑、疟疾合适的预警界值为 P_{80} ,麻疹、肺结核合适的预警界值为 P_{90} 。

【关键词】 传染病;控制图法;预警界值

A study on early detection for seven infectious diseases YANG Wei-zhong*, XING Hui-xian, WANG Han-zhang, LAN Ya-jia, SUN Qiao, HU Shi-xiong, LV Wei, YUAN Zheng-an, CHEN Yu-xu, DONG Bai-qing. *Chinese Center for Disease Control and Prevention, Beijing 100050, China

【Abstract】 Objective To provide methods and alert thresholds which are scientific, sensitive, specific and practical for Early Warning System in Public Health Surveillance. **Methods** Alert data was based on historical infectious diseases reports. Control chart was used to detect outbreaks or epidemics. An epidemic was defined by consulting Specialists. After calculating sensitivity, specificity, positive predictive value and describing receiver-operating characteristic curve(ROC), the optimal model and thresholds were chosen. **Results** At 80 percentile, the sensitivities and the specificities of epidemic haemorrhagic fever, hepatitis A, dysentery, epidemic cerebrospinal meningitis and malaria were over 90%, and there was a high efficacy of early warning. At 90 percentile, the sensitivities and the specificities of tuberculosis and measles were over 85%, and there was a high efficacy of early warning also. **Conclusion** Control chart based on five years was chosen as a essential method in early warning system. The alert threshold for epidemic haemorrhagic fever, hepatitis A, dysentery, epidemic cerebrospinal meningitis and malaria was 80 percentile. The alert threshold for tuberculosis and measles was 90 percentile.

【Key words】 Notifiable infectious disease; Control chart; Alert threshold

开展对传染病疫情的监测预警工作是疾病监测信息体系的重要内容,我国虽然较早建立了传染病疫情监测报告系统,但由于缺少有效的分析手段,未很好地对历史资料进行开发,以致到现在仍未实现对传染病疫情进行预警的工作,很难早期发现传染病的爆发和流行,使传染病的防治工作在一定程度上处于被动应付的局面。2003 年春季发生的 SARS 疫情,使人们更加认识到建立传染病早期预警系统的迫切性和必要性。

资料与方法

1. 研究病种:肾综合征出血热(出血热)、甲型肝炎(甲肝)、细菌性痢疾(痢疾)、流行性脑脊髓膜炎(流脑)、疟疾、麻疹、肺结核 7 种传染病。

2. 数据库来源:以广西壮族自治区、湖南省、辽宁省、福建省和上海市共 366 个县 1997-2003 年 7 种传染病月报数据建立数据库。366 个县 1998-2003 年甲乙类传染病报告发病率分布与全国同期甲乙类传染病报告发病率分布相同。

3. 确定流行标准:经典的“流行”定义指某种疾病的发病水平超过历年一般水平,这是一个定性的概念,实际应用时需要量化。本研究采用专家咨询法对“流行”这个定义进行量化,以此作为判断历史数据中疾病流行的参照标准。

基金项目:科技部科研院所社会公益研究专项基金资助项目(2002DIA40020)

作者单位:100050 北京,中国疾病预防控制中心(杨维中、邢慧娟、王汉章);四川大学公共卫生学院(兰亚佳);上海市疾病预防控制中心(孙乔、袁政安);湖南省疾病预防控制中心(胡世雄、陈裕旭);广西壮族自治区疾病预防控制中心(吕伟、董柏青)

4. 建立预警模型:按照控制图法预警模型原理^[1],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,计算出指定的百分位数(如 P_{50} 、 P_{60} 、 P_{65} 、 P_{70} 、 P_{75} 、 P_{80} 、 P_{90} 、 P_{95}),作为候选的预警界值,建立预警模型(图 1)。

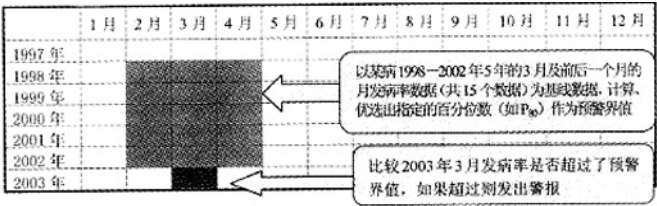


图1 5 年基线数据建立预警模型的原理

5. 优选合适的预警界值:用 1997-2002 年研究病种的月发病率数据作为基线数据,代入模型,用各候选预警界值对 2003 年的传染病疫情分县分病分月进行预警,同时用流行参照标准对其分县分病分月判断是否流行,从而计算各候选预警界值预警的灵敏度、特异度、阳性预测值等指标,并绘制 ROC 曲线(receiver-operating characteristic curve),比较各候选预警界值预警功效,从中优选出合适的预警界值(图 1)。

6. 统计软件:采用 EXCEL、SPSS、FOXPRO。

结 果

1. 流行标准:将调查表寄送给 50 名专家,其中 38 名为医学院校教授及疾病预防控制中心具有丰

富实践经验的流行病学专家,上述专家参加了全部三轮咨询调查,调查表回收率 76%。对所研究的 7 种传染病的调查结果显示:51.3%~71.8% 的专家认为“流行”定义中“历年”以 5 年为好;52.6%~59.0% 的专家认为,以大于“ $\bar{x} + 2s$ ”判断疾病流行,即“超过一般水平”。

2. 合适的预警界值:在综合平衡灵敏度、特异度、阳性预测值、预测功效以及预测疾病的特点后,优选出 7 种传染病的合适预警界值:麻疹、肺结核为第 90 百分位数;出血热、甲肝、痢疾、流脑、疟疾为第 80 百分位数。预警出血热、痢疾、甲肝、流脑、麻疹和疟疾时的灵敏度和特异度均在 90% 以上,预警肺结核时灵敏度和特异度分别为 88.1% 和 85.9%,所有 7 种疾病的阳性预测值在 44.1%~79.2% 之间(表 1 和图 2)。

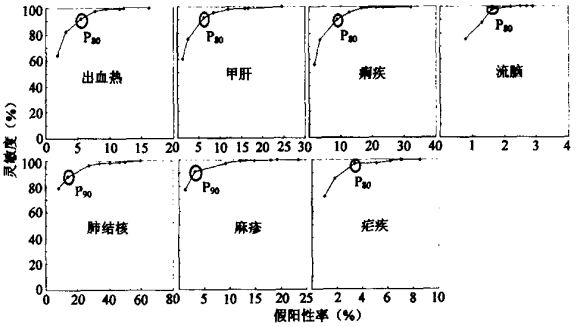


图2 7 种传染病预警的 ROC 曲线(5 年控制图法)

表1 7 种传染病各候选预警界值的灵敏度(%)、特异度(%)、阳性预测值(%)

疾病	指标	P_{95}	P_{90}	P_{80}	P_{75}	P_{70}	P_{65}	P_{60}	P_{50}
出血热	灵敏度	63.9	82.4	91.6	97.5	99.2	99.2	100.0	100.0
	特异度	98.2	96.8	94.5	92.3	90.6	88.5	87.8	83.8
	阳性预测值	66.7	59.4	48.7	41.9	37.5	32.8	31.7	25.9
甲肝	灵敏度	60.2	75.8	91.8	95.8	98.1	98.9	98.9	100.0
	特异度	98.9	97.5	93.8	91.5	88.1	84.2	83.4	75.5
	阳性预测值	78.3	67.1	49.7	42.9	35.6	29.5	28.5	21.5
痢疾	灵敏度	56.3	75.1	90.0	95.8	98.7	99.0	99.0	99.7
	特异度	98.4	96.5	91.1	87.3	83.4	78.8	76.5	67.9
	阳性预测值	73.7	62.5	44.1	37.0	31.7	26.7	24.8	19.5
麻疹	灵敏度	77.0	91.3	97.6	99.2	99.7	100.0	100.0	100.0
	特异度	98.7	96.8	90.8	88.0	85.2	82.1	80.8	76.7
	阳性预测值	88.8	79.2	59.0	52.8	47.7	42.9	41.3	36.7
流脑	灵敏度	75.6	87.8	97.5	100.0	100.0	100.0	100.0	100.0
	特异度	99.2	98.7	98.4	97.8	97.5	97.3	97.3	97.1
	阳性预测值	75.6	69.2	66.7	59.4	56.2	53.9	53.9	52.6
肺结核	灵敏度	79.1	88.0	96.4	98.0	98.5	99.0	99.3	100.0
	特异度	91.9	85.9	72.7	66.2	59.7	53.9	50.1	40.2
	阳性预测值	72.2	62.4	48.5	43.6	39.5	36.4	34.6	30.8
疟疾	灵敏度	71.4	85.7	97.4	97.4	98.7	100.0	100.0	100.0
	特异度	99.0	98.2	96.6	94.9	93.8	93.0	92.9	91.5
	阳性预测值	78.6	71.0	59.1	49.3	45.0	42.3	41.8	37.4

讨 论

预警方法的灵敏度、特异度和阳性预测值是通过比较传染病预警系统发出警报的情况与流行参照标准判定“流行”的情况而得到的^[2,3]。但是,如何判定“流行”,目前国际上没有统一的定量标准^[4,5]。在发病率已经极低的一些国家,有以发病数的绝对数作为判断“流行”的标准^[6,7],而本研究所涉及的疾病在我国是常见高发的传染病,所以,采用绝对数作为判断“流行”的标准显然不合适。因此,本研究选用专家咨询法来确定一个用于比较的“流行”参照标准。

评价预警方法的灵敏性和准确性的基本指标是灵敏度、特异度和阳性预测值,在优选预警界值时,一个重要原则是取得这三项指标的平衡。ROC 曲线可以很直观的显示出这个较好的平衡点^[2],即曲线上最靠近坐标轴左上角的那点,以此点代表的界值作为系统的预警界值,预警的功效可能最大。但在实际工作中,合适界值的选取既要考虑预警的功效,同时又要考虑不同疾病的特点,或偏重于灵敏度或偏重于特异度来选择界值。在选出血热的合适预警界值时,如单纯从 ROC 曲线看,最靠近坐标轴左上角的点是 P_{90} ,但是根据出血热的发病特点,预警时更偏重于高的灵敏度。 P_{80} 点的灵敏度比 P_{90} 上升了 9.2%,而特异度仅下降了 2.3%,所以选择 P_{80} 作为合适预警界值更佳。同理,疟疾的预警界值没有选择最靠近坐标轴左上角的 P_{90} ,而是选择了 P_{80} 作为合适预警界值。

选择较长时间的基线数据可能会提高预警的准确性,但传染病流行趋势的变化会增加基线数据的变异,从而抵消了长基线带来的好处^[8]。自 1990 年以后,我国多数传染病呈显著下降趋势,近 5、6 年来,下降趋势趋于平缓,比较稳定的维持在一个较低水平(肺结核除外)。在这种情况下,基线数据时间太长,会大大降低预警的敏感性。所以,我们选择 5 年作为基线数据时间。用 5 年基线数据建模,数据数量偏少,一些统计指标的计算和模型的稳定有可能受到影响。为解决这种矛盾,美国疾病预防控制中心作预警时,将预警时间单位向前和向后各移动一个单位,数据量增加了 2 倍,证明这种处理方法是可行的^[4,9]。本研究借用了这个方法,将观察月既往 5 年该月及其前后一个月的月发病率数据都作为基线数据,使原来的 5 个数据增加为 15 个数据。

受我国传染病报告历史报告的限制,本研究所

建立的预警数据库是以传染病月报数据为基础的,与国外报道的周预警方法比较^[10,11],要预警一些短潜伏期的传染病爆发可能存在时间滞后性缺陷,解决这个问题有赖于今后积累足够长时间的周报告数据(至少 5 年)。

本研究确定的预警技术主要面向县级单位,如果预警覆盖区域过大,模型不能对局部地区出现的疫情做出灵敏的预警^[12]。如果预警覆盖区域过小,比如村、镇,由于发病数据不稳定,可能无法建立这种预警模型。

本研究确定的控制图法预警技术,具有简单易行、适用面广、预警功效高的特点,避免了象其他预测方法那样需要不断地对参数进行调整的弊端。在现行的传染病网络实时直报的条件下,基线数据会随着传染病日常报告数据的更新而自动更新,从而保证预警系统持续处于工作状态。

(广西壮族自治区疾病预防控制中心、湖南省疾病预防控制中心、辽宁省疾病预防控制中心、福建省疾病预防控制中心和上海市疾病预防控制中心为本研究提供了传染病疫情报告数据,谨致感谢)

参 考 文 献

- 1 章扬熙.医学统计预测.北京:中国科技出版社,1995.
- 2 Michael MW, Tsui FC, Jeremy UE. The emerging science of very early detection of disease outbreaks. J Public Health Management Practice, 2001, 7: 51-59.
- 3 Framework for evaluating public health surveillance systems for early detection of outbreak: recommendations from the CDC working group. MMWR, 2004, 53 RR05: 1-11.
- 4 Costagliola D, Flahault A, Galinec D, et al. A routine tool for detection and assessment of epidemics of influenza-like syndromes in France. Am J Public Health, 1991, 81: 97-99.
- 5 Snacken R, Lion J, Casteren V, et al. Five years of sentinel surveillance of acute respiratory infection (1985-1990): the benefits of an influenza early warning system. Eur J Epidemiol, 1992, 8: 485-490.
- 6 Melinda W, William P. Evaluation of a method for detecting outbreaks of diseases in six states. Am J Prevent Med, 1993, 9: 45-49.
- 7 José G Rigau-Pérez, Peter SM. A deviation bar chart for detecting dengue outbreaks in Puerto Rico. Am J Public Health, 1999, 89: 374-378.
- 8 Donna FS, Melinda W. Evaluation of a method for detecting aberrations in public health surveillance data. Am J Epidemiol, 1993, 137: 373-380.
- 9 Changes in format for presentation of notifiable disease report data. MMWR, 1990, 39: 234.
- 10 Najera J, Kouznetsov R, Dilacollette C. Malaria epidemics. Detection and control. Forecasting and prevention. Geneva, World Health Organization, 1998.
- 11 Early warning systems. Communicable disease surveillance & response (CSR). 2004. <http://www.who.int/csr/labepidemiology/projects/earlywarnsystem>
- 12 Shuji H, Yoshitaka M. Detection of epidemics in their early stage through infectious disease surveillance. Inter J Epidemiol, 2000, 29: 905-910.

(收稿日期:2004-08-10)

(本文编辑:张林东)