

Bayes概率预测法在控制八十年代流行性脑脊髓膜炎流行高峰中的应用

曾光¹ 胡真¹ 杨天英² 李林村² 来秀君⁸ 吴贵坤³ 陈长志⁴ 魏建军⁵
 朱伯健⁶ 王福泉⁷ 阎亚林⁸ 李金城⁶ Stephen B. Thacker⁹ 吴泰顺¹
 徐昌¹ 徐涛¹ 柔克明¹ 朱宝兰² 李瑞兰² 徐长松² 王全庆⁴ 杜世君⁴
 张清彦⁴ 郭爱芝⁵ 刘福贵⁵ 王相义⁵ 轩勤生⁶ 李培珍⁶ 赵文立⁶ 刘红敏⁶
 俞文祥⁷ 刘建民⁸ 任守礼⁸ 郑善平⁸ 温宏伟⁸

摘要 在八十年代流行性脑脊髓膜炎(简称流脑)流行高峰到来之前,本协作组应用Bayes概率法或原理相似的优度法在北京、河南、江苏等地以不同行政区划的方式开展了流脑流行早期数理预测的研究。结果表明本研究可在流脑流行高峰的半年之前提出预测值,并将预测结果用于指导菌苗分配。经多年观察,预测区的流脑发病率下降幅度明显大于非预测区。

关键词 流行性脑脊髓膜炎 预测 数学模型 贝叶斯定理

流行性脑脊髓膜炎(简称流脑)在我国是一种发病率高、危害严重的急性传染病。但在八十年代以前我国流脑防制工作基本处于被动防制局面,既缺乏特异性的防制措施更缺乏预测指导下的防制决策。1980年我国研制的流脑多糖菌苗试用成功,但当时限于菌苗生产能力及各级防疫机构财力有限,不能开展普种。为使有限的菌苗以及人力物力财力获得最大的防制效果,控制可能在八十年代出现的流脑流行高峰,1982年由中国医学科学院流研所(现属中国预防医学科学院)牵头成立了流脑预测协作组,开展了Bayes概率法预测流脑流行的研究。1982年至今,已在国内一亿人口以上流行区应用本法进行了实际预测,取得了显著的社会和经济效益。

预测的统计学基础

通过对各地流脑疫情资料进行认真观察研究发现,上年度流行时间分布曲线的偏度、跨度、曲线下的面积等与下年度流脑发病率的高低以及升降趋势之间存在一定的统计学联系。

偏度可用病例分布的10%、20%、50%(中位数)、80%和90%分位数以及众数旬(月)次来表示,曲线下面积可用年发病率高低来表示。

经过各地资料实际考核验证,不同地区的同一指标大多具有同向性规律。以病例累计分布的10%分位数旬(月)次为例,则河南10个地区(1960~1981年),北京10个远郊县区(1960~1980年),郑州市7个县区(1964~1983年)和江苏盐城市8个县区(1958~1985年)的流脑疫情资料一致显示出,如上年旬(月)次数值偏小,则下一年流脑发病率高的概率小,反之,如上一年数值偏大,则下一年流脑发病率高的概率大。

1 中国预防医学科学院流行病学微生物学研究所

2 河南省卫生防疫站

3 北京市卫生防疫站

4 河南省唐河县卫生防疫站

5 河南省郑州市卫生防疫站

6 河南省商丘地区卫生防疫站

7 江苏省盐城市卫生防疫站

8 河南省许昌市卫生防疫站

9 美国疾病控制中心

Bayes概率法(和优度法)可将不同指标反映出的概率差异综合在一起,作出一定效能的判别分析,从而达到预测的目的。

我们对这些统计学联系存在的原因有如下解释:这些现象很可能反映了流脑的严格季节性与人群免疫水平消长之间复杂的关系。例如,上一年流行高峰后偏,则很可能说明在流行后期仍然存在着较多的易感人群和具备流脑流行的其它条件,只是由于气候变暖终止了流行,这些条件可能延续到下一年。因此,下一年发生流行或发病率升高的危险性要加大。反之,如上一年流行高峰前偏,则在流行后期仍然具备流行的气候条件下病例已不多了,这很可能因为易感者已不多或者其他流行条件不具备所致,意味着在下一个流行年度发生流行或者发病增多的机会也不多。

资料与方法

一、资料来源与整理:所用材料为各预测研究地区流脑疫情、人口统计、菌苗预防接种、流脑带菌调查以及人口额外流动量等资料,各地区选取的资料大同小异。

各研究地区历年流脑流行期优势菌群都是A群,各地区流行规律相似,本地区内流脑病例的诊断标准及病例报告方法相同。

要按照流脑流行周期重新划分流行年度。规定自上年11月至当年10月为一个流行周期,其中上年11月至下年7月为流行期,8~10月为间歇期。间歇期病例数一般都少于1%,因此只用流行期资料作分析。按旬(月)整理流行期资料,依次划定1~27旬次或1~9月月次,计算不同的分位数或众数的旬、月次和中间60%或80%分位数旬(月)期。

二、变量选择与规定:

1.两个因变量:

流行强度(Y):可分为流行、散发两类,或大流行、流行、小流行散发四类。各地依本地发病率高低酌情规定划分界值。

流行趋势(Z):分为升高年(环比值≥

150%)和非升高年(环比值<150%)。

2.16个初选自变量: X_1 (上年流行期流行强度), X_2 (上年流行趋势), X_3 (上年流行期病例发生的众数旬/月次), X_5 、 X_6 、 X_7 和 X_8 (依次为上年流行期病例累计发生的10%、20%、80%和90%分位数旬/月次), X_9 、 X_{10} (分别为上年度流行期80%、60%病例发生的旬/月期), X_{11} (流行后期病例数占全期病例数之比), X_{12} (为上一年流行年度人群流动量), X_{13} (本年度人群流动量), X_{14} (人群带菌率), X_{15} (上年流行后期发病率)和 X_{16} (预防接种因素*)。

各地可根据本地区的实际情况适当选择,重新编号。

三、计算方法: Bayes 概率法的计算原理及计算过程同文献^[1,2]。亦可采用原理一致算法更简便的优度法进行计算^[3]。

四、对比算法: 北京、河南唐河县卫生防疫站分别应用逐步判别分析法^[4]做为对比算法。

五、模型检验方法: 在使用模型前除常规作复测检验外,为检验模型的外推能力,普遍开展了回顾性预测检验。

1. 回顾性预测检验之一——样本分段检验^[5]: 将手头资料分为两段,保留后段样本(一般为2~3年)不纳入建模样本,仅以前段样本为分析基础建立数学模型,以后段样本作外推检验。由于后段样本包括了近期信息,检验效果好有助于对实际预测(即继续外推)的信心。

2. 回顾性预测检验之二——随机抽样检验: 从总样本中随机抽出部分样本(10%左右)不参加建模,在用剩余样本建模后,用抽出的样本作外推性检验。由于按随机原则抽出的样本包括了各种不同水平的流行规律和流行趋势,可较全面地检测模型的外推能力。

六、预防接种后预测方法的调整: 上述计

*根据接种率和保护率调整预测结果时用

算方法是建立在对流脑流行无干预基础上的，而实施流脑菌苗接种是预防流脑流行的特异性措施，大面积地接种必然打破流脑自然流行的规律，使得实际发病率下降，增加预测难度。因此在预测研究中，我们重视了预防接种因素对预测的影响，按照预防接种率的高低对预测结果予以调整。

根据文献记载，我国生产的流脑多糖菌苗免疫后三年间的保护率分别为97.64%、79.56%和67.07%^[7]。在调整预测结果时以此为参考依据，根据当年和既往两年接种率高低给以加减分，通过对各地流脑流行强度和流行趋势数据拟合试验，建立了因变量两类水平调整方法和多类水平调整方法。

七、预防接种后预测结果的现场考核：预测结果为流行或升高，而采取菌苗接种措施后，未出现流行或升高的结果，那么如何验证原来的预测是否正确呢？

我们采取了以下两种方式予以考核：

1. 随机选取对照区：预测1985年河南唐河县要发生流脑流行。该县共有13个乡，因菌苗

数量有限，不能对全部易感者进行接种，故随机选7个乡作常规菌苗接种。对其它6个乡不做常规接种，一旦发现爆发点，采取应急包围接种。这样做的结果，可近似地观察到原预测结果是否准确。除唐河县外，1984年河南通许县也从12个乡中随机抽取1个乡作应急接种对照区。

2. 将未接种的地区或人群作对照：例如，观察1985年商丘、驻马店和南阳地区因各种原因未接种菌苗的县、乡以此作为对照区来考核原预测结果。

结果与分析

一、模型检验结果：为检查流脑预测模型的预测效果，采用了全样本复测和随机复测以及抽样回顾性和分段回顾性预测来检验模型的拟合度。各地符合率分别为80~85%、80~100%、75~85%和75~100%。

二、模型实际预测结果：为检验模型的实际预测能力，各地分别用当地资料建立的模型进行实际预测（表1）。

表1

我国部分地区流脑预测模型的预测符合率(%)

单 位	年 份	预 测 数	Bayes概率法		优 度 法	
			流 行 强 度	流 行 趋 势	流 行 强 度	流 行 趋 势
北京市卫生防疫站	1983	10	100.00	70.00	—	—
河南省卫生防疫站	1985~86	18	—	—	100.00	100.00
唐河县卫生防疫站	1984~89	50	—	—	93.30	88.33
商丘地区卫生防疫站	1986	5	100.00	100.00	—	—
郑州市卫生防疫站	1987	7	100.00	100.00	100.00	85.71
许昌市卫生防疫站	1988	5	—	—	100.00 (100.00)	80.00 (80.00)
江苏盐城市卫生防疫站	1986~88	24	100.00	100.00	—	—

三、对比方法计算结果：为比较Bayes概率法、优度法与其它预测方法的优劣。北京市和河南唐河县用逐步判别分析法，分别进行了模型预测研究。结果Bayes概率法与优度法的预测结果与对比算法相似，但由于计算方法简单，比对比算法更适宜在基层推广。

四、预测效果的人群实验研究结果：见表

2和表3。

流脑预测与流脑控制

一、指导流脑防治：利用流脑预测模型预测流脑的流行动态，掌握流脑发生流行的时间、地区和强度，为制定流脑防治策略和措施提供科学依据，进而为卫生行政部门决策提供科学

表2 河南省流脑预测流行地区菌苗接种人群
实验研究结果

地区	年份	分组	乡数	易感人口	发病人数	发病率(/10万)
通许县	1984	接 种	11	412747	17	4.12
		未接 种	1	33022	17	51.48
唐河县	1985	常规接种	7	394950	48	12.15
		应急接种	6	317030	85	26.81

表3 河南预测1985年流脑发生流行地区菌苗接种和未接种组的发病情况

地区	分组	易感人口	发病数	发病率(/10万)
商丘	未接种	139006	91	65.46*
	接 种	1560471	39	2.50
驻马店	未接种	381032	426	111.80*
	接 种	1518389	98	6.45
南 阳	未接种	698651	909	130.11*
	接 种	1710928	130	7.60

*未接种组发病率如此之高，说明流行高峰的确已经来临

依据。

河南省及所属各地（市）北京市以及江苏省盐城市都依据预测结果采取了有针对性的菌苗接种措施。

以河南省为例，该省原属流脑高发省份，自1966年后约10年出现一次流行高峰。1976~1977年流脑大流行后，1982年起发病率再度上升，1984年全省发病34575例，发病率达45.55/10万，而且预测1985年疫情将继续上升，将预测结果上报省卫生厅后，引起了高度重视，卫生厅及时成立全省流脑防治领导小组，召开全省防治工作会议，决定增购600万人份流脑菌苗，并拨10万元专款。商丘、南阳、驻马店、周口等地区行署或卫生局得知本

地预测结果后，也增加了几十万元流脑防治专款使这些地区1985年的流脑疫情不但没有上升，反而出现了大幅度下降。反之，在预测不流行的年份和地区，则没有进行菌苗普种，避免了不必要的投资，使有限的卫生资金得以充分的利用。

二、开展预测后的流脑防制效果：流脑预测地区自1985年以来，应用Bayes概率法或优度法预测流脑的发病动态，从而指导流脑防制取得了显著效果，流脑预测地区的流脑发病率下降速度明显高于全国的发病率（表4）。

表4 1984年以来流脑预测地区和全国流脑发病率的比较

地 区	1984~1988年流脑发病率(/10万)				
	1984	1985	1986	1987	1988
河南省	45.55	11.55	3.35	1.63	1.57
商丘地区	106.25	6.40	1.66	0.70	0.81
唐河县	33.61	18.24	3.74	1.54	1.26
许昌市	45.01	10.03	3.02	1.19	1.21
郑州市	36.01	12.22	3.79	1.57	1.12
北京市	3.88	1.99	1.53	1.34	0.77
全 国	11.69	10.73	7.56	3.21	1.97

三、社会效益与经济效益：在预测指导下合理使用菌苗，取得巨大的经济效益与社会效益。按全国同期资料估算，1985~1988年期间，仅河南省减少发病79795例，减少病死4388人，合计减少经济损失2190万元（表5）。

河南省与毗邻省比较，河南省减少发病60547人，减少死亡3330人，减少经济损失1647万元。八十年代辽宁、江苏和河南三省使用菌苗较多，每个易感者（1~15岁）平均接种菌苗2.98、2.08和1.83次，辽宁、江苏比河南分别多1.15和0.25次，由于河南在Bayes概率法预测结果指导下合理使用菌苗，八十年代发病率下降幅度大于辽宁、江苏省。

表5 1984~1988年河南省减少流脑发病率(/10万)和发病数估算表

年份	全国		河南省发病率			河南省人口数	减少发病人数
	发病率	实际值	估算值	减少值			
1984	8.99	45.55					
1985	10.67	11.55	54.06	42.51	76794804	32645	
1986	7.89	3.55	39.98	36.43	77603503	28271	
1987	3.34	1.63	16.92	15.29	78706567	12034	
1988	2.00	1.57	10.13	8.58	79777025	6845	
合计							79795

注：合计减少死亡数=79795×5.5%*=4388(人)

合计减少经济损失=79795×157.6**+4388×2080**=21702732(元)

* 5.5%为全国流脑平均病死率

** 据文献[7]记载，每减少1例流脑发病可减少157.6元，每减少1例死亡可节约2080元

The Use of Bayes' Theorem in Controlling the Coming Epidemic Peak of Epidemic Meningitis in the 1980s Zeng Guang, et al., Institute of Epidemiology and Microbiology, Chinese Academy of Preventive Medicine, Beijing

To control the coming epidemic peak of epidemic cerebrospinal meningitis (ECM) in the 1980s in the regions with 100 million population in China, the mathematical models based upon Bayes' theorem (BT) were established and used respectively in provincial, regional and county's level. Reports of ECM from each ten-day's period or each month during the meningitis season were analysed to create forecast models. Records of ECM vaccinating rate in previous years were fully taken into account to modify the theoretical values. Calibration, split-sample, random-sample selection, as well as actual forecast tests, were used to check the efficiency of the models. The distribution of meningitis vaccine was planned according to the final predictive results. The incidence rates of ECM of above regions decreased obviously faster than other areas in China. Attributing the application of BT forecast research, it was estimated only in Henan

province 79795 ECM cases; 4388 deaths and 21 million Yuan economic damage were avoided during the 4 years period, from 1985 to 1988.

Key words Epidemic cerebrospinal meningitis (ECM) Forecasting Mathematical model Bayes' theorem.

参 考 文 献

- Guang Zeng, et al. An assessment of the use of Bayes' theorem for forecasting in public health: the case of epidemic meningitis in China. Int J Epidemiol 1988; 17 (3) : 673.
- 曾光, 等. 应用Bayes概率法早期预测流脑的流行. 中华流行病学杂志 1985; 6 (6) : 347.
- 陈长志, 等. 应用尤度法进行县以下流脑流行早期数理预测方法的研究. 中华流行病学杂志 1985; 6 (6) : 384.
- 上海第二医学院数学教研室. 判别分析法在医学上的应用(内部资料), 1980.
- Armstrong JS. Long-range forecasting. New York: Johnwiley and Sons, 1978: 303, 312.
- 魏荣泰, 等. 流脑提纯菌苗人体免疫反应及连续三年效果观察. 中华流行病学杂志 1982; 3 (1) : 16.
- 李金城, 等. 应用流脑多糖菌苗控制流脑大流行及经济效益评价(内部资料). 1987.

(1989年9月10日来稿, 1989年10月8日修回)