张姝丽，孙培源，张德山，等. 2009. 北京地区细菌性痢疾医疗气象预报模式及流行风险水平评估［J］. 气候与环境研究，14（4）：673－677. Zhang Shuli，Sun Peiyuan，Zhang Deshan，et al. 2009. The basic research of medical meteorology forecast model of bacillary dysentery and its risk valuation in Beijing area［J］. Climatic and Environmental Research (in Chinese)，14（4）：673－677.

北京地区细菌性痢疾医疗气象预报模式及流行风险水平评估

张姝丽[1] 孙培源[2] 张德山[1] 蒲永兰[2] 韦懿云[2] 赵娜[1] 宛公展[3]

1 北京市专业气象台，北京 100089
2 北京市海淀区疾病预防控制中心，北京 100086
3 天津市气象科学研究所，天津 300081

摘 要 研究了北京市海淀区2004～2006年全年逐日细菌性痢疾发病率与地面气象要素的相关关系。结果发现，细菌性痢疾发病与水汽压（空气绝对湿度）高度相关，单相关系数平均可达0.85以上。在所有候选预报因子中，它的方差贡献可占80%。采用多元回归概率分级技术，建立了细菌性痢疾医疗气象预报及风险水平评估模式。该模式预报结果可为各级疾病预防控制中心指导广大城乡居民卫生防疫提供科学依据。

关键词 细菌性痢疾 疾病流行风险水平评估 医疗气象预报
文章编号 1006-9585（2009）06-0673-05 **中图分类号** P49 **文献标识码** A

The Basic Research of Medical Meteorology Forecast Model of Bacillary Dysentery and Its Risk Valuation in Beijing Area

ZHANG Shuli[1], SUN Peiyuan[2], ZHANG Deshan[1], PU Yonglan[2], WEI Yiyun[2], ZHAO Na[1], and WAN Gongzhan[3]

1 *Beijing Municipal Weather Information Service, Beijing* 100089
2 *Beijing Haidian Control Centre of the Disease Prevention , Beijing* 100086
3 *Tianjin Research Institute of Meteorology, Tianjin* 300081

Abstract The relationship between the grounded meteorological elements and the daily incidence of bacillary dysentery in Haidian District of Beijing from 2004 to 2006 has been studied. The result shows that the incidence of bacillary dysentery is highly related with the vapor pressure (absolute humlidity), and the simple correlation coefficient is over 0.85 averagely. Compared with all the selected elements, the contribution of variance can reach 80%. The forecast of bacillary dysentery and the mode of risk identification are established by using the classification technology of multivariate regression. The forecast model can provide scientific basis for the centers of disease control and prevention at all levels, which are in charge of sanitation and epidmic prevention for the urban and rural residents.

Key words bacillary dysentery, risk valuation of disease, medical meteorology forecast

收稿日期 2008-05-19 收到，2009-04-30 收到修定稿
作者简介 张姝丽，女，1955年出生，高级工程师，主要从事专业气象预报服务工作。E-mail：shuli333@sohu.com

1 引言

细菌性痢疾，是由痢疾杆菌引起的急性肠道传染病，由于患病后免疫力不能持久，且菌型繁多，加之耐药菌株的不断出现，其发生和流行仍较严重（谈荣梅等，2003）。目前，细菌性痢疾在我国的流行仅次于流行性感冒，也是北京地区夏秋季节的常见流行性疾病或疫情，其发生发展和蔓延与环境气象条件密切相关。根据已有的历史病例和同期天气气候资料，研究细菌性痢疾与天气气候的相关关系，建立医疗气象预报模型，可逐日滚动提供天气气候与疾病疫情发生的趋势预报与风险水平评估，在很大程度上为疾病预防控制中心提供疾病可能流行的气候背景，有利于贯彻“预防为主”的方针，指导广大城乡居民防病治病，保障居民身体健康（张德山等，1999）。

2 资料搜集与因子相关分析

2.1 资料来源

细菌性痢疾逐日发病病例，取自北京市海淀区疾病预防控制中心。2004年1月1日～2006年12月31日共搜集到病例样本1092个，有4天缺测。病例涵盖人口约为250万。同步地面气象因子由北京市观象台提供，包括9个逐日地面基本气象要素（平均气温、极端最高气温、极端最低气温、平均相对湿度、最小相对湿度、平均水汽压、日降水量、平均风速、平均海平面气压）。因子样本数9×1096个，没有缺测。

2.2 细菌性痢疾疾病发病的统计学特征

细菌性痢疾，一年四季都有发生，但主要发病期在5～10月，季节性十分明显。盛夏时发病率达到最高，说明此期间的温、压、湿、风等气象要素的环境最适宜痢疾杆菌的繁殖和生存（见图1）。

表1为细菌性痢疾发病的统计参数值。从表1不难看出，细菌性痢疾日平均发病人次每百万人口约为6.2，中位数为3.7（把样本数据依大小排序后，位于中间的数为中位数。当样本数目为奇数时，最中间的数就是中位数；当样本数目为偶数时，中间的两个样本数据的算术平均数为中位数）。由于中位数远小于平均值，概率分布显然不属于正态分布类型。

2.3 细菌性痢疾发病与环境气象因子的相关分析

所谓环境气象因子，是指根据9个地面基本气象要素按一定方法加工派生而成（屠其璞等，1984；宛公展，1988）。其中，因子1、4、5、6、7、8、9是因子源值的等权或半衰权重的平均值。因子2、10变化倾向值的计算方法是：首先将因子

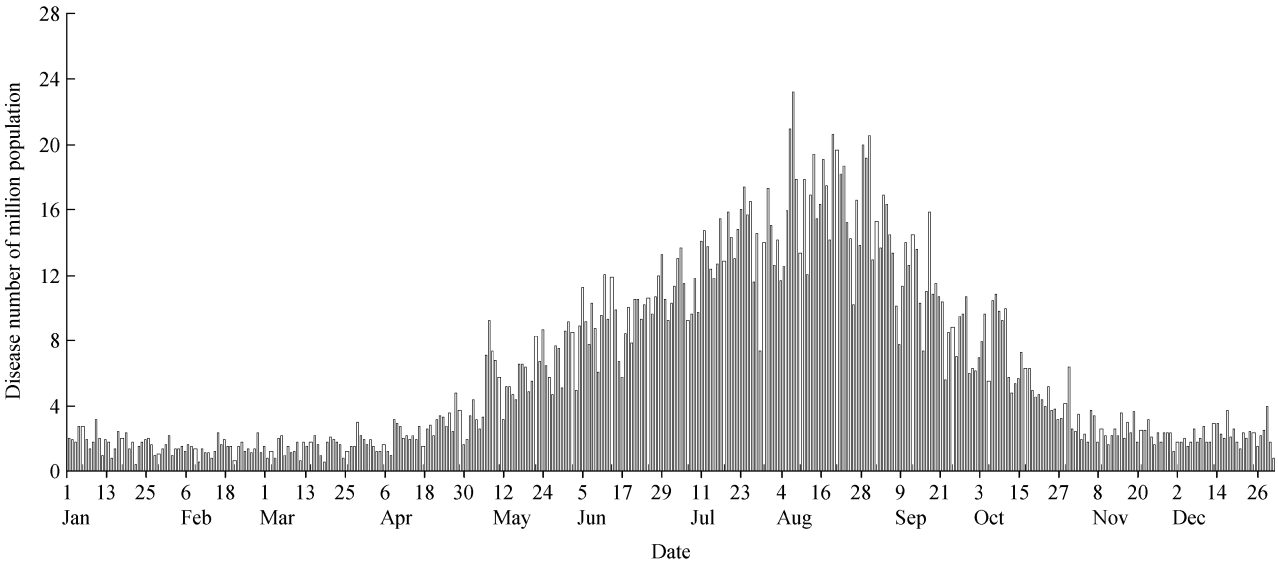


图1 2004～2006年细菌性痢疾发病平均季节变化

Fig. 1 Histogram of average season variety about the germs dysentery from 2004 to 2006

表 1　北京市海淀区细菌性痢疾逐日发病的统计特征值
Table 1　Daily statistic characteristic of bacillary dysentery incidence in Haidian District of Beijing

特征分类	统计参数	出现日期
最大值	32.0	2005-08-18
次大值	31.6	2005-08-22
第 3 大值	31.2	2005-08-15
最小值	0.4	2004-01-04
中位数	3.7	
平均值	6.2	
标准差	5.9	

注：统计参数为人次/每百万人口

源值用直线方程拟合，再将解得的斜率值乘以方程相关系数的绝对值。为了找出细菌性痢疾发病与环境气象因子之间的相关关系，采用经典方差贡献分析技术，现将部分计算结果列于表 2。因子 X_i 的方差贡献值 PV_i，即因子的偏回归平方和，是因子重要性的度量方法之一，常用下式计算，

$$PV_i = L_{yy}^{-1} B_i^2 C_{ii}^{-1}, \tag{1}$$

其中，B_i^2 为因子 X_i 的偏回归平方和，L_{yy} 为细菌性痢疾发病的原方差，C_{ii} 为正规方程系数逆矩阵对角线元素。为便于不同因子相互比较，表 2 中将其处理为百分比（表中的方差贡献值、总体复相关系数、统计量 F 值均是预测时效 24～168 h 的参数）。

表 2　细菌性痢疾发病与气象因子的相关分析
Table 2　The correlation analysis between bacillary dysentery incidence and meteorological factors

环境气象因子	方差贡献值	单相关系数
日平均气温	0.1%	0.7882
气温变化倾向	0.2%	−0.0165
气温日较差	3.6%	−0.3271
平均相对湿度	1.1%	0.6920
最小相对湿度	0.7%	0.6860
平均水汽压	89.7%	0.9065
日降水量	0.2%	0.5266
平均风速	0.2%	−0.3989
平均海平面气压	4.2%	−0.6664
气压变化倾向	0.0%	0.0555
总体复相关系数		0.9493
统计量 F 值		949.41

从表 2 可见，细菌性痢疾的发病与表示空气湿度和温度的水汽压及日平均气温相关密切，单相关系数分别达到 0.9065 和 0.7882（当样本数目超过 1000 时，通过信度 $\alpha = 0.01$ 检验的相关系数下限阈值 R_α 仅为 0.081），水汽压的方差贡献值可占总贡献的 89.7%。方差贡献较大的因子还有平均海平面气压、气温日较差等，其贡献与上述因子贡献之和可达 97%。其次，所有因子的复相关系数平均可达 0.9493，统计量 F 值可达 949.41。远远超过了通过 F 检验所需的下限阈值（$\alpha = 0.05$，$F_\alpha = 1.83$）。证明，细菌性痢疾发病与气象因子群高度相关，利用环境气象因子的变化预测北京地区未来细菌性痢疾发病趋势，是可行的。

3　医疗气象预报建模与评估疾病流行风险水平的思路与方法

细菌性痢疾医疗气象预报与流行风险水平评估采用 5 档分级的方法。不同级别对应不同疾病发病人数及流行风险水平。为此，通过以下两个步骤完成。

3.1　组建逐日发病人次滚动预报的多元回归模式

本文采用经典回归技术建立预报模型。该模式可写为

$$Y_i = B_\sigma + \sum B_j \times X_j, \tag{2}$$

式中，Y_i 为每百万人口的平均日发病人次的回归拟合值。X_j 为进入方程的预报因子。B_σ（$j = 1, 2, \cdots, m$）为相应的回归系数，可通过求解正规方程获取。经典回归模式的正规方程为

$$\boldsymbol{X}_X \boldsymbol{B} = \boldsymbol{Y}_X \tag{3}$$

式中，$\boldsymbol{X}_X = \boldsymbol{X}'\boldsymbol{X}$ 为因子协方差矩阵，可以根据因子方差贡献的大小预选 6、7、10 或 11 个等不同的因子组合搭配。$\boldsymbol{Y}_X = \boldsymbol{X}'\boldsymbol{Y}$，为相应气象因子与细菌性痢疾发病率的交叉平方和，$\boldsymbol{B}$ 为回归系数向量。

具体建模之前，需从已经搜集的原始样本序列中，连续滑动截取预报单元（$P+T$），作为对应不同预报时效的有效样本备用。前者对应过去 P 天的气象因子连续取值区间，后者对应未来 T 天平均发病的取值区间。不同建模时效对应的各年最大可能样本数为：$365 + L_p -（P + T）+ 1$

(闰年 $L_p=1$，否则为0)。

考虑到痢疾杆菌从适宜的温湿环境中生长繁殖到通过苍蝇等传播媒介污染生冷食物，最终感染人体，且在人体内有1～3天的潜伏期，即：从适宜的气象条件到细菌性痢疾疾病疫情发生存在滞后性，分别用前3天、前5天和前7天（取 $P=3$、5、7）的气象因子建模，预测未来1～7天（$T=7$）的发病人数，比较其预测拟合情况。

结果表明，前7天的气象因子所建模型拟合率优于前3天和前5天的，误差±1级的拟合率达到98%以上，而且拟合率随着预报时效的延伸而提高（见图2）。

因此，取 $P=7$，求解（4）式得到细菌性痢疾的6个时效回归参数，给出使用10个预报因子预测细菌性痢疾病的未来24～168 h 的回归模式：

$$Y_{A168} = -78.4737 + 0.0137b_1 + 0.1397b_2 - 0.1159b_3 - 0.0444b_4 - 0.0402b_5 + 0.7911b_6 - 0.0531b_7 - 0.1574b_8 + 0.0809b_9 + 0.0289b_{10}, \quad (4)$$

式中的 Y_{A168} 为细菌性痢疾的未来24～168 h 百万人口发病人次的预测值，b_1、b_2、b_3、…、b_{10} 依次为表2中的10个气象因子。

3.2 疾病预报分级与流行风险水平的评估

应用概率积分的方法进行同步评估疾病流行风险水平与预报结果的气象分级，要点是：

首先，按照评估疾病流行风险水平（或气象分级）的实际需要，确定不同的分级概率标准。比如，概率标准定为5级，各级所占概率分别为

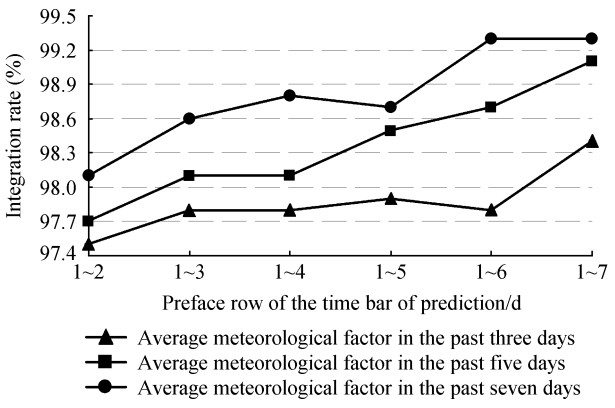


图2 三种时效气象因子的模式预测细菌性痢疾发病的拟合率（≤±1级）

Fig. 2 Fitting rates (≤±1 class) of bacillary dysentery prediction in the forcast system using three leading time scales factors

表3 细菌性痢疾疾病医疗气象预报概率分级

Table 3 Probability classifications of the medical meteorology forecast about bacillary dysentery

预报概率	预报级别	发病人数	流行风险水平
$Y_i>g_4$	5	显著增多	很高
$g_3<Y_i\leqslant g_4$	4	较多	较高
$g_2<Y_i\leqslant g_3$	3	呈现增多趋势	中等
$g_1<Y_i\leqslant g_2$	2	较少	较低
$Y_i<g_1$	1	很少	很低

P_5、P_4、P_3、P_2和 P_1。注意，$\sum P_i=1$（$i=5$，4，…，1），这将意味着（3）式的预报结论中，将有 $P_i\times100\%$的预报结果可以期望降落到第 i 级。

其次，如何判断（2）式预报结果 Y_i 的级别归属，如果仍将概率标准定为5级，在样本的连续域上为了分级须有4个分点即 g_4、g_3、g_2和 g_1（见表3）。

分界指标 g_1、g_2、g_3和 g_4的获取，可以根据指定的概率密度标准 P_1、P_2、P_3、P_4、P_5 对（2）式作概率积分，当曲边梯形的面积分别达到 $\sum P_i$时，所记录的4个分点位置 g_i（$i=1$，2，3，4）即为5级的分界指标。

4 结果分析

4.1 所建模型对历史资料的回代模拟

首先，指定一种预报时效及相应的概率分级标准，对所有历史样本实况进行同步分级（含流行风险水平评估，下同），再用同期气象因子值回代（3）式，获得所有建模样本的拟合预报分级。现以概率密度标准 0.15、0.20、0.30、0.20、0.15为例（过去7天气象资料，预报未来24～168 h），拟合结果的比较见表4。

表4 不同预报时效历史样本拟合率的比较

Table 4 Comparisons between different fitting rates of prediction and the historical samples in different leading times

预测时效/h	样本数	完全正确	相差±1级	正确率（含±1级）	失误率
24～48	1069	54.764%	40.241%	95.005%	4.995%
24～72	1066	58.163%	37.384%	95.547%	4.453%
24～96	1063	59.163%	37.302%	96.465%	3.535%
24～120	1060	59.701%	37.687%	97.388%	2.612%
24～144	1057	59.963%	37.605%	97.568%	2.432%
24～168	1054	60.788%	36.773%	97.561%	2.439%

从表 4 数据可以看出，如果以预报结果与实况相差±1 级以内算作正确的话，不同时效的拟合正确率，平均可达 96%以上。预报级别与实况完全一致的，在 54%～60%之间。随着预报时效的增加，平均拟合准确率略有增加。

4.2　不同概率密度标准对拟合准确率的可能影响

在 5 级分档概率密度标准中，如果适当降低两端概率密度 P_1、P_5的取值，即调整到医学常用的概率密度分布函数：0.03594，0.23831，0.4515，0.23831，0.03594。注意，需保证 $\sum P_i = 1$。仍使用过去 7 天气象资料，预报未来 2～7 天的模型，则拟合准确率的变化如表 5。

由表可见总的拟合正确率平均提高了 1%～2%，级别完全符合平均可提高 10%。经过这样的调整，拟合准确率有明显改善，但预报产品结果中，出现在 1 级或 5 级的概率必然有所降低，这与细菌性痢疾流行风险很大是小概率事件的现象相吻合。

5　服务应用检验

应用细菌性痢疾疾病预测模式，预测出未来 6 种时效逐日的细菌性痢疾发病等级与北京市海淀区 2007 年逐日监测病例比较，统计出的预报准确率是理想的（见表 6）。

6　结语

综上所述，细菌性痢疾的发病流行与前期的气象条件存在较高的相关关系，从各气象要素与发病数的单相关系数可见，潮湿、高温、低气压是适宜痢疾杆菌生长繁殖的气象条件，其中与水气压关系最为密切，细菌性痢疾发病的预报准确率较为理想，各预报时效的完全正确率平均为 70.49%，误差±1 级的正确率平均达到 99.67%。

细菌性痢疾的发病除了适宜的前期气象条件外，还与传播媒介、个人卫生习惯等因素相关联。因此，本研究的工作尚属初步，还需与相关部门开展深入的合作，不断完善细菌性痢疾的发病的预测模式。

表 5　调整概率密度标准后不同预报时效历史样本拟合率
Table 5　The fitting rates in different leading times after adjusting the probability density

预测时效/h	样本数	完全正确	相差±1 级	正确率（含±1 级）	失误率
24～48	1069	63.922%	33.580%	97.502%	2.498%
24～72	1066	65.306%	32.375%	97.681%	2.319%
24～96	1063	67.070%	30.698%	97.767%	2.233%
24～120	1060	67.164%	30.690%	97.854%	2.146%
24～144	1057	67.540%	30.309%	97.848%	2.152%
24～168	1054	67.917%	30.488%	98.405%	1.595%

表 6　北京市海淀区 2007 年细菌性痢疾预测试验准确率
Table 6　Accuracy rate of bacillary dysentery prediction for Haidian District in the test period of 2007

预报时效/h	样本数	完全正确	误差±1 级	误差≤1 级	误差>1 级
24～48	356	67.98%	31.46%	99.44%	0.56%
24～72	355	71.55%	28.17%	99.72%	0.28%
24～96	354	69.77%	29.66%	99.44%	0.57%
24～120	353	69.69%	30.03%	99.72%	0.28%
24～144	352	71.88%	27.84%	99.72%	0.28%
24～168	351	72.08%	27.92%	100.00%	0.00%
平均		70.49%		99.67%	

参考文献（References）

谈荣梅，陈坤，屠春雨. 2003. 气象因素与细菌性痢疾发病关系的探讨［J］. 浙江预防医学，15（3）：7－9. Tan Rongmei, Chen Kun, Tu Chunyu. 2003. Study on the relationship between meteorological factors and incidence of bacillary dysentery［J］. Zhejiang Journal of Preventive Medicine (in Chinese), 15 (3): 7－9.

屠其璞，王俊德，丁裕国，等. 1984. 气象应用概率统计学［M］. 北京：气象出版社，540pp. Tu Qipu, Wang Junde, Ding Yuguo, et al. 1984. Meteorologic Application Probability and Statistics［M］ (in Chinese). Beijing: China Meteorological Press, 540pp.

宛公展. 1988. PC－1500 袖珍计算机 BASIC 语言应用程序集［M］. 北京：气象出版社，132. Wan Gongzhan. 1988. PC－1500 Mini Computer Basic Language Program Collection［M］(in Chinese). Beijing: China Meteorological Press, 132.

张德山，王玉玲，谢庄. 1999. 天气变化影响人体健康的信息与预报［J］. 气候与环境研究，4（1）：32－34. Zhang Deshan, Wang Yuling, Xie Zhuang. 1999. The information and forecasting of impact on human health by weather variation［J］. Climate and Environmental Research (in Chinese), 4 (1): 32－34.