

# 安徽省2004—2006年疟疾环境影响因素研究

王丽萍 方立群 许娴 王建军 马家奇 曹务春 金水高

**【摘要】** 目的 研究安徽省2004—2006年疟疾传播的相关环境影响因素。方法 采用空间分析技术建立安徽省乡镇尺度疟疾疫情环境影响因素空间数据库,综合应用主成分分析和logistic回归分析方法对数据进行统计学处理。结果 2004—2006年安徽省疟疾疫情呈现地区聚集性,淮北地区疫情逐年加重。乡镇尺度疟疾疫情主要与当地温度、降雨量、归一化差值植被指数和海拔等因素有关。“年最低气温(温度系列指标的4主成分)”每升高1个单位,疟疾发生的可能性减少33%;“年降雨总量(降雨量系列指标的第1主成分)”每增加1个单位,疟疾发生的可能性将减少27%;“海拔”每升高10 m,疟疾发生的可能性减少2%;而归一化差值植被指数与疟疾发生可能性的关系与前3个因素不同,每增加1个单位,疟疾发生的可能性将增加3.28倍。结论 2004—2006年安徽省淮北地区是该省2000年以来疟疾疫情回升后出现的新时空聚集区,淮北地区的地形地貌、温度、降雨量等自然环境因素影响其疟疾发病水平,这将为进一步探索研究2000年后以安徽省为代表的中国中部地区疟疾回升的主要原因提供重要参考。

**【关键词】** 疟疾; 发病率; 主成分分析; logistic回归分析

**Study on the determinants regarding malaria epidemics in Anhui province during 2004—2006** WANG Li-ping\*, FANG Li-qun, XU Xian, WANG Jian-jun, MA Jia-qi, CAO Wu-chun, JIN Shui-gao. \* National Center for Public Health Surveillance and Information Service, Chinese Center for Disease Control and Prevention, Beijing 100050, China

Corresponding author: JIN Shui-gao, Email: jinsg@vip.sina.com

**【Abstract】 Objective** To analyze the determinants of malaria in Anhui province during the year of 2004—2006. **Methods** A GIS-based database of malaria at the township scale of Anhui province was established, using remote sensing and spatial analysis technologies. Methods on statistical analysis, principal component analysis combined with logistic regression analysis were synthetically used to analyze the association between malaria and environmental factors. **Results** Malaria epidemics in Anhui province during 2004—2006 mainly occurred in the northern districts of Huai River, and the epidemics had become more serious yearly. The determinants of malaria at the township scale mainly included factors as temperature, rainfall, normalized difference vegetable index and elevation. If the lowest temperature in a year, which appeared the fourth principal component of the temperature index series, increased by one unit, the probability of incidence of malaria would decrease by 33%. If the total annual rainfall, which was the first principal component of the rainfall index series, increased by one unit, the probability would decrease by 27%. If the elevation increased by 10 meters, the probability would decrease by 2%. However, the relationship between Normalized Difference Vegetable Index (NDVI) and the probability of incidence of malaria was different. If the NDVI had a one unit increase, the probability would increase 3.28 times. **Conclusion** The northern districts of Huai River during 2004—2006 appeared to be a new spatio-temporal cluster when reemergence of malaria epidemics had occurred in Anhui province since 2000. Terrain and physiognomy, nature and circumstances factors, such as temperature and rainfall had affected the incidence rates of malaria. Our research data from Anhui province would provide some important references to the discovery of main reasons on the reemergence of malaria epidemics since 2000, especially in the central geographic areas of China.

**【Key words】** Malaria; Incidence rate; Principal component analysis; Logistic regression analysis

20世纪50—60年代,我国疟疾流行严重,发病率在100/万左右,70年代以后大幅度下降,80—90年代

初,疫情虽有小幅调整,但均未形成流行高峰,20世纪末全国疟疾疫情下降到0.3/万左右<sup>[1]</sup>。2000年后我国中部地区以皖北为中心的沿淮河流域地区疟疾发病大幅回升,连续多年出现局部暴发。研究表明1990—2006年间安徽省疟疾疫情向北部移动,2004—2006年的安徽省北部地区成为该省疟疾疫情新的时空聚集区<sup>[2]</sup>。影响疟疾流行的因素复杂,寄生

DOI: 10.3760/cma.j.issn.0254-6450.2009.01.011

基金项目:国家自然科学基金资助项目子项目(30590371)

作者单位:100050 北京,中国疾病预防控制中心公共卫生监测与信息服务(王丽萍、马家奇、金水高);军事医学科学院微生物流行病学研究所(方立群、曹务春);安徽省疾病预防控制中心(许娴、王建军)

通信作者:金水高,Email:jinsg@vip.sina.com

虫、蚊媒、人类宿主和环境等因素及其交互作用决定了疟疾传播和感染的风险<sup>[3]</sup>。近年来地理信息系统(GIS)结合卫星遥感技术及全球定位系统已成为研究空间邻近度及传染病环境、气象因素变化的最主要手段<sup>[4]</sup>。本研究利用2004—2006年安徽省疟疾疫情监测资料,并收集整理相关的环境气象资料,通过空间分析技术与统计学相结合的方法,研究疟疾疫情的主要环境影响因素。

## 资料与方法

1. 疫情数据:来自2004—2006年安徽省《疾病监测信息报告管理系统》各乡镇疟疾疫情发病统计数据。

2. 乡镇级数字区划图:应用安徽省1:5万地形图,几何校正、投影后逐一矢量化乡镇界限,录入乡镇名称,建立乡镇界限图层,由中科永生数据公司协助完成。

3. 环境气象数据获取方法:

(1)气温、降雨量:据国家气象局提供的安徽省及周边省(江苏、山东、河南、湖北、江西、浙江)的112个气象站点数据。数据提取方法:①应用克里金法逐月生成安徽省栅格化的气象图;②然后在ArcGIS 9.0软件中将生成的各月份气温、降雨量地图与乡镇界线地图进行叠加,应用空间分析方法提取各乡镇月均气温及降雨量。

(2)归一化差值植被指数(normalized difference vegetable index, NDVI)、湿度因子及水体:通过解译专题制图仪(thematic mapper, TM)遥感数据获取。具体方法:①中国科学院遥感应用研究所提供安徽省2005—2006年TM遥感影像12景,应用安徽省1:5万地形图对每景TM数据进行几何校正;②通

过遥感图像处理软件(environmental for visualizing images, ENVI)提取每景TM数据的NDVI值,并将生成的12景NDVI图片进行合并,生成安徽省NDVI地图;③同时对每景TM数据进行K-T变换(缨帽变化),提取其第三主成分湿度因子(wetness),合并后生成安徽省湿度因子地图;④水体的解译在每景TM数据叠加安徽省2000年土地利用地图的基础上,应用监督分类的方法提取每景TM数据的水体,并将其合并后转化为矢量格式,生成安徽省矢量化水体图层;⑤将上述生成的NDVI、湿度因子及水体图层与安徽省乡镇界线图进行叠加,应用空间分析方法分别计算各乡镇的平均NDVI、平均湿度因子及水体面积。

4. 统计学分析:采用主成分分析方法,对温度系列指标和降雨量系列指标进行主成分分析,确定相应的主成分。之后采用经典的多元logistic回归分析方法对乡镇尺度疟疾疫情影响因素进行研究,因变量为2004—2006年间各乡镇是否有疟疾病例报告,自变量因素包括了相应的温度(主成分)、降雨量(主成分)、NDVI、湿度指数、水体及GDP等因素,对模型进行检查基础上最终建立logistic回归模型。

## 结果

1. 安徽省各乡镇疟疾发病率分布:根据患者所属的乡镇与安徽省1:5万乡镇界线数字地图建立空间关联,绘制各年度乡镇级疟疾发病率空间分布专题地图(图1)。可以看出,3年间安徽省北部乡镇疟疾发病率升高明显,疫情逐年加重,相对而言安徽省南部乡镇发病水平较低,出现明显的地区聚集性特点。

2. 乡镇尺度疟疾疫情影响因素:

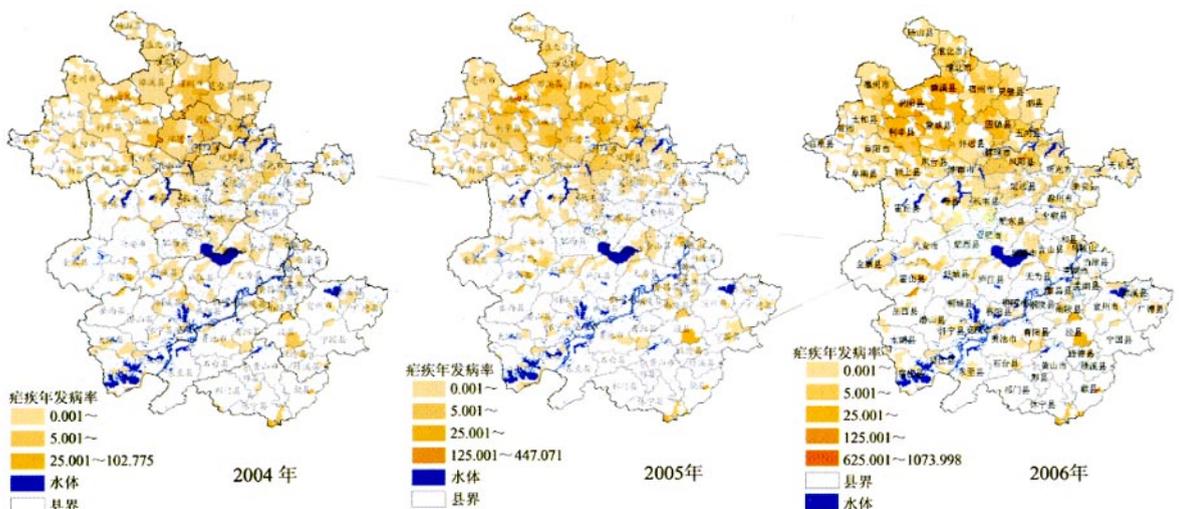


图1 2004—2006年安徽省各乡镇疟疾发病率空间分布

(1) 主成分分析:16个温度指标的主成分分析结果见表1、2。C1~C4的贡献率分别为67.18%、20.50%、7.72%和2.37%,累计贡献率达97.77%,即利用了原资料约98%的信息。从特征向量来看,第一特征向量的各分量之大小大致相当,说明第一特征向量是综合指标,同时可以看出在年平均气温上有较大系数,说明反映的是年平均气温;第二特征向量在7月份气温和年最高气温上有较大的系数,说明反映的是夏季/年最高气温;第三特征向量在3—4月份气温上有较大系数,说明反映的是春季气温;第四特征向量在1月份和年最低气温上有较大系数,说明反映的是冬季/年最低气温。

表1 温度指标的16个特征根

| 序号 | 特征根     | 差值     | 贡献     | 累计贡献   |
|----|---------|--------|--------|--------|
| 1  | 10.7489 | 7.4695 | 0.6718 | 0.6718 |
| 2  | 3.2793  | 2.0449 | 0.2050 | 0.8768 |
| 3  | 1.2345  | 0.8547 | 0.0772 | 0.9539 |
| 4  | 0.3798  | 0.2580 | 0.0237 | 0.9777 |
| 5  | 0.1217  | 0.0232 | 0.0076 | 0.9853 |
| 6  | 0.0985  | 0.0408 | 0.0062 | 0.9914 |
| 7  | 0.0577  | 0.0136 | 0.0036 | 0.9950 |
| 8  | 0.0442  | 0.0220 | 0.0028 | 0.9978 |
| 9  | 0.0222  | 0.0156 | 0.0014 | 0.9992 |
| 10 | 0.0066  | 0.0039 | 0.0004 | 0.9996 |
| 11 | 0.0027  | 0.0002 | 0.0002 | 0.9998 |
| 12 | 0.0025  | 0.0016 | 0.0002 | 0.9999 |
| 13 | 0.0009  | 0.0003 | 0.0001 | 1.0000 |
| 14 | 0.0006  | 0.0006 | 0.0000 | 1.0000 |
| 15 | 0.0000  | 0.0000 | 0.0000 | 1.0000 |
| 16 | 0.0000  | 0.0000 | 0.0000 | 1.0000 |

表2 温度指标的前4个特征向量

| 变量    | C1      | C2      | C3      | C4      |
|-------|---------|---------|---------|---------|
| 月均温度  |         |         |         |         |
| 1     | 0.2713  | -0.1805 | -0.0052 | 0.5000  |
| 2     | 0.2848  | -0.1681 | 0.0898  | -0.0084 |
| 3     | 0.2416  | -0.1417 | 0.4622  | -0.1490 |
| 4     | 0.2511  | -0.0480 | 0.4731  | -0.2833 |
| 5     | 0.2855  | -0.0031 | 0.2607  | -0.0715 |
| 6     | 0.1372  | 0.4357  | 0.3266  | 0.1277  |
| 7     | 0.1230  | 0.4934  | -0.0085 | 0.2567  |
| 8     | 0.2680  | 0.1866  | -0.1911 | -0.2243 |
| 9     | 0.2804  | 0.1605  | -0.1860 | -0.1938 |
| 10    | 0.2772  | 0.1141  | -0.2171 | -0.3458 |
| 11    | 0.2651  | -0.0859 | -0.3812 | -0.0494 |
| 12    | 0.2755  | -0.0425 | -0.3383 | -0.0712 |
| 年最低气温 | 0.2713  | -0.1805 | -0.0052 | 0.5000  |
| 年最高气温 | 0.1257  | 0.4915  | -0.0127 | 0.2551  |
| 年平均气温 | 0.3048  | 0.0005  | -0.0109 | -0.0562 |
| 年温度变异 | -0.2296 | 0.3558  | 0.0381  | -0.1641 |

13个降雨量指标的主成分分析结果见表3、4。C1~C4贡献率分别为65.65%、16.65%、9.28%和4.13%,累计贡献率达95.71%,即利用了原资料约96%的信息。从特征向量来看,第一特征向量的各分量之大小大致相当,说明第一特征向量是综合指标,同时可以看出在年降雨总量上有较大系数,说明反映的是年降雨总量;第二、三特征向量在9月、8月降雨量上有较大的系数,说明反映的是夏末初秋的降雨量;第四特征向量在12月降雨量上有较大系数,说

表3 降雨量指标的13个特征根

| 序号 | 特征根    | 差值     | 贡献     | 累计贡献   |
|----|--------|--------|--------|--------|
| 1  | 8.5351 | 6.3706 | 0.6565 | 0.6565 |
| 2  | 2.1645 | 0.9578 | 0.1665 | 0.8230 |
| 3  | 1.2067 | 0.6696 | 0.0928 | 0.9159 |
| 4  | 0.5371 | 0.3178 | 0.0413 | 0.9572 |
| 5  | 0.2193 | 0.0812 | 0.0169 | 0.9741 |
| 6  | 0.1381 | 0.0625 | 0.0106 | 0.9847 |
| 7  | 0.0756 | 0.0140 | 0.0058 | 0.9905 |
| 8  | 0.0616 | 0.0259 | 0.0047 | 0.9952 |
| 9  | 0.0357 | 0.0156 | 0.0027 | 0.9980 |
| 10 | 0.0201 | 0.0138 | 0.0015 | 0.9995 |
| 11 | 0.0063 | 0.0063 | 0.0005 | 1.0000 |
| 12 | 0.0000 | 0.0000 | 0.0000 | 1.0000 |
| 13 | 0.0000 | 0.0000 | 0.0000 | 1.0000 |

表4 降雨量指标的前4个特征向量

| 变量    | rC1     | rC2     | rC3     | rC4     |
|-------|---------|---------|---------|---------|
| 月降雨量  |         |         |         |         |
| 1     | 0.3264  | -0.1363 | -0.0254 | -0.1975 |
| 2     | 0.3033  | 0.1854  | 0.3083  | 0.0659  |
| 3     | 0.2344  | -0.2772 | -0.3108 | 0.5547  |
| 4     | 0.3189  | 0.1563  | 0.1507  | -0.1434 |
| 5     | 0.3080  | 0.2456  | 0.1846  | 0.0657  |
| 6     | 0.3265  | -0.1361 | -0.0254 | -0.1976 |
| 7     | -0.2575 | 0.1833  | 0.4805  | 0.2894  |
| 8     | 0.0208  | -0.4833 | 0.6044  | 0.2205  |
| 9     | -0.1312 | 0.6099  | 0.0060  | 0.0594  |
| 10    | -0.3125 | 0.1561  | 0.0350  | 0.0553  |
| 11    | 0.3179  | 0.1872  | -0.0570 | -0.2333 |
| 12    | 0.2490  | 0.2219  | -0.2469 | 0.6231  |
| 年降雨总量 | 0.3141  | 0.1385  | 0.2977  | 0.0586  |

明反映的是入冬后降雨量。

(2) 多元logistic回归分析:多元logistic分析变量说明见表5。以“2004—2006年乡镇是否有疟疾报告”为因变量,温度、降雨量、NDVI、湿度指数、海拔、年人均GDP、乡镇水体面积及构成等为解释变量构建结局变量为“二分类变量(CASE\_YN)”的多元logistic回归模型。逐步logistic多元分析结果见表6,最终的模型可以表达为:

$$\text{logit}P = -0.456 - 0.398(\text{冬季/年最低气温}) - 0.315(\text{年降雨总量}) - 0.021(\text{海拔}) + 1.455\text{NDVI}$$

以 $\alpha$ 水准为0.05,“冬季/年最低气温”、“年降雨总量”、“海拔”和“NDVI”4个因素为安徽省乡镇疟疾发生的主要影响因素。在控制模型中其他自变量不变的情况下,“冬季/年最低气温(温度系列指标的第4主成分,无量纲)”每升高1个单位,疟疾发生的可能性减少33%;“年降雨总量(降雨量系列指标的

表5 多元logistic回归分析变量说明

| 序号 | 变量               | 单位含义说明               | 变量名      |
|----|------------------|----------------------|----------|
| 1  | 年人均GDP           | 1=500元               | Gdp_500  |
| 2  | 海拔               | 1=10 m               | Dem_10   |
| 3  | C1~C4温度主成分       | 无量纲                  | -        |
| 4  | C1~C4降雨量主成分      | 无量纲                  | -        |
| 5  | NDVI             | 无单位                  | Ndvi     |
| 6  | 温度因子             | 无单位,1=50             | Wet_50   |
| 7  | 水体面积             | 1=500 m <sup>2</sup> | Water_50 |
| 8  | 水体构成             | 1=5%                 | Wpro_5   |
| 9  | 2004—2006年疟疾病例报告 | 1=有,0=无              | Case_yn  |

表6 安徽乡镇疟疾发病影响因素 logistic 回归模型分析

| 模型变量    | $\beta$ | $s_e$ | z 值     | P 值     | OR 值 (95%CI)        |
|---------|---------|-------|---------|---------|---------------------|
| 常数项     | -0.456  | 0.128 | 12.653  | 0.000   | -                   |
| 冬季年最低气温 | -0.398  | 0.120 | 10.909  | 0.001   | 0.672(0.531~0.851)  |
| 年降雨总量   | -0.315  | 0.025 | 162.778 | <0.0001 | 0.730(0.695~0.766)  |
| 海拔      | -0.021  | 0.007 | 9.537   | 0.002   | 0.979(0.967~0.992)  |
| NDVI    | 1.455   | 0.715 | 4.135   | 0.042   | 4.283(1.054~17.406) |

第1主成分,无量纲)“每增加1个单位,疟疾发生的可能性将减少27%”“海拔”每升高10 m,疟疾发生的可能性减少2%;而NDVI与疟疾发生可能性的关系与前3个因素不同,即NDVI每增加1个单位,疟疾发生的可能性增加3.28倍。

## 讨 论

安徽省位于暖温带与亚热带的过渡带,气候多变,复杂的地形地貌、气候特点、南北地形及气候上的差异为蚊虫等传染病传播媒介的孳生提供了有利或不利的因素,这些为理解安徽省内疟疾发病的南北差异提供了参考。

本研究所采用的疟疾疫情监测数据及环境影响因素数据来源权威。由于我国目前没有现成的乡镇级气温及降雨量数据,文中通过空间分析方法进行提取。为了减少环境数据提取偏性,在站点的选择上,综合考虑安徽省及6个周边省份112个气象站点监测数据。同样在遥感影像的处理上,也充分注意包括覆盖安徽省边角的遥感影像图,使提取的数据具有极强的代表性。

研究表明,在排除传染源数量、人群免疫状况及防控措施影响因素情况下,从疟疾传播危险性指标来看,淮北地区的地形地貌、气象条件等因素是疟疾流行水平的主要影响因素。海拔升高将使疟疾发病风险降低。淮北地区与皖南山区相比,海拔较低,因此发生疟疾的风险也相应增加。该研究结果与非洲等疟疾高流行区的结果类似<sup>[5-7]</sup>。安徽省年降雨量在南北地区也存在明显差异,年平均降雨量由北向南为770~1770 mm,分析中发现降雨量增加可以降低疟疾发病的风险,安徽省南部地区年降雨量高于淮北地区,其疟疾发生风险则低于淮北地区,也成为解释淮北地区疟疾高发的一个因素。事实上,研究表明,降雨量对疟疾流行的影响较为复杂,一般随地区与媒介蚊种孳生习性的不同而有显著差异,降雨量还通过影响环境的湿度而影响疟疾发病水平<sup>[8,9]</sup>。温度是疟疾传播一个非常重要的指标,在拟合模型的“年平均气温”、“夏季/年最高气温”、“春季气温”和“冬季/年最低气温”4个主成分中,仅“冬季/年最低

气温”对乡镇间疟疾发病率的差异有关系,且分析结果显示乡镇“冬季/年最低气温”每升高1个单位,其发生疟疾的可能性减少33%,与类似的研究结论有所不同。大部分研究结果显示,某一地区温度升高将使疟疾传播的风险增加<sup>[10-12]</sup>。笔者考虑,不排除地区横向分析造成的偏性。最后一个显著的影响因素是NDVI。NDVI可综合反映蚊虫孳生环境植被覆盖、温度、湿度、降雨量等综合生态信息,大量的疟疾遥感研究表明NDVI是反映蚊虫孳生地生态信息的良好指标<sup>[13]</sup>。本研究与众多研究结果相似,发现NDVI每增加1个单位,发生疟疾的可能性增加3.28倍。

本研究结果将为进一步探讨2000年后我国以安徽省为代表的中部地区疟疾疫情回升的主要原因提供依据。研究中所采用的温度、降雨量及相关遥感计算指标数据来源权威,空间分析提取方法科学,目前在其他类似研究中,其尺度大多为县区级,尚未有乡镇级研究结果的报道。与县区级研究尺度相比,将研究尺度进一步缩小至乡镇,可以将县区所辖各乡镇的各项指标值凸显出来,即研究精度进一步加大,将更有利于探测这些指标在地域上的变化,从而使因素分析结果更具有现实性和可解释性。文中将空间分析技术与统计分析方法综合应用于疟疾研究领域,可为类似研究提供方法学参考。

## 参 考 文 献

- [1] 周水森,王漪,汤林华. 2005年全国疟疾形势. 中国寄生虫学与寄生虫病杂志, 2006, 24(6): 401-403.
- [2] 王丽萍,徐友富,王建军,等. 1990-2006年安徽疟疾疫情时空分布特点研究. 疾病控制杂志, 2008, 12(2): 156-159.
- [3] 朱继民,汤林华. 疟疾及其蚊媒的预测方法研究进展. 中国寄生虫学与寄生虫病杂志, 2007, 25(2): 150-153.
- [4] 王丽萍,金水高. GIS空间分析技术在疟疾研究中的应用. 中国公共卫生, 2008, 24(6): 745-747.
- [5] Hartman JEK, McConnell K, Chan N, et al. Climate suitability for stable malaria transmission in Zimbabwe under different climate change scenarios. Glob Change Hum Health, 2002, 3: 42-54.
- [6] Taylor P. The malaria problem in Zimbabwe epidemiology. Cent Afr J Med, 1985, 31(9): 163-165.
- [7] Taylor PMS. A review of the malaria situation in Zimbabwe with a special reference to the period 1972-1984. Trans R Soc Trop Med Hyg, 1986, 80: 12-19.
- [8] 温亮,徐德忠,张治英. 基于GIS的疟疾预警系统构建的初步分析. 中国公共卫生, 2004, 20: 626-627.
- [9] 贾尚春,邹铮,徐伏牛. 全球气候变暖对疟疾传播的潜在影响. 中国寄生虫病防治杂志, 2004, 17: 63-65.
- [10] Sérandour J, Girel J, Boyer S, et al. How human practices have affected vector-borne diseases in the past: a study of malaria transmission in Alpine valleys. Mala J, 2007, 6: 115-125.
- [11] Hunter PR. Climate change and waterborne and vector-borne disease. J Appl Microbiol, 2003, 94 Suppl: S37-46.
- [12] 刘亦仁. 对我国中部地区疟疾流行与控制的几点认识. 中国寄生虫病防治杂志, 2005, 18: 241-243.
- [13] 于国伟,汤林华. 遥感地理信息系统技术在疟疾研究中的应用. 中国寄生虫学与寄生虫病防治杂志, 2004, 22: 119-121.

(收稿日期:2008-09-16)

(本文编辑:张林东)