

# 江苏省肝癌与地理因素回归分析

覃 玉<sup>1</sup>, 胡晓抒<sup>1</sup>, 赵金扣<sup>1</sup>, 陆应昶<sup>1</sup>, 武 鸣<sup>1</sup>, 王培桦<sup>1</sup>, 周晓农<sup>2</sup>, 杨国静<sup>3</sup>

(1.江苏省疾病预防控制中心, 江苏 南京 210009;

2.中国疾病预防控制中心寄生虫防治所, 上海 200032;

3.江苏省血吸虫病防治所, 江苏 无锡 214000)

**摘 要:** [目的]运用地理信息系统 (geographic information system, GIS) 探讨江苏省肝癌与地理因素之间的流行病学关系。 [方法]用 ArcView 3.0a 软件建立江苏省肝癌流行病学 GIS 数据库, 从 USGS (U.S Geological Survey) 地理数据库中提取我省各地区地理数据, 在 ArcView 3.0a 软件支持下用 SPSS 软件对肝癌和各项地理指标作回归分析, 进行最佳曲线拟合。 [结果]肝癌的高发地区局限于江苏省中部的沿海地区。肝癌死亡率随海拔高度的增加而降低, 而与温度、降雨量、地面蒸发量呈显著正相关。 [结论]肝癌死亡率的分布与地理因素相关; 地理信息系统在癌症流行病学中的运用需做进一步的探讨。

**关键词:** 肝肿瘤; 地理信息系统; 地理; 流行病学; 江苏

中图分类号: R73-31; R735.7

文献标识码: A

文章编号: 1004-0242(2003)11-0630-03

## Regression Analysis on Liver Cancer and Geographic Factors in Jiangsu Province

QIN Yu, HU Xiao-shu, ZHAO Jin-kou, et al.

(Jiangsu Provincial Center for Disease Control and Prevention, Nanjing 210009, China)

**Abstract:** [Purpose] To explore the correlation of liver cancer and geographic factors by geographic information systems (GIS). [Methods] The ArcView 3.0a software was used for setting up GIS epidemic databases of liver cancer in Jiangsu province. Geographic factors on each point was gotten from USGS (U.S Geological Survey). Regression analysis on liver cancer and geographic factors was analyzed by SPSS at the basis of ArcView 3.0a to fit the best regression curve. [Results] Coastal areas in medial Jiangsu province were the high-risk areas for liver cancer. Liver cancer mortality increased with the decreasing of altitude, while the mortality had positive correlation with temperature, rainfall and evaporation. [Conclusions] The distribution of liver cancer mortality is related to the geographic factors. GIS can be applied to cancer epidemiology, which should be further explored.

**Key words:** liver neoplasms; geographic information system (GIS); geography; epidemiology; Jiangsu

肝癌是严重危害人民生命健康的疾病之一, 1990年至1992年江苏省人口全死因调查资料显示, 肝癌的死亡率为36.81/10万, 居全省恶性肿瘤死亡的第一位<sup>[1]</sup>。以往对疾病分布及流行的描述多局限于某一个或几个地区的资料, 并且缺少直观性。近年来, 地理信息系统 (geographic information system, GIS) 技术的成熟和发展, 为流行病学研究提供了新

的分析工具, 目前GIS已成为很多公共卫生和流行病学项目的一个重要组成部分。本文将借助GIS对肝癌与地理因素之间的关系作流行病学分析。

## 1 材料与方法

### 1.1 材 料

基础地图: 江苏省县界电子地图 (1:100万) 由国

收稿日期: 2003-06-16; 修回日期: 2003-08-18

家基础地理信息中心提供,并与疾病数据库相联获分县的GIS数据库备用。

肝癌数据库的建立:收集我省开展肿瘤登记地区1996年~1998年肝癌死亡资料,包括年死亡数、年人口数,建立肝癌死亡数据库。

地理资料数据库的建立:从USGS(U.S Geological Survey)地理数据库中提取我省各地区30年(1940年~1970年)的点地理数据,各地理点距离为15公里,包括海拔高度,年平均温度,年平均降雨量,月平均降雨量,年平均蒸发量,月平均蒸发量,建成江苏省地理资料数据库。

## 1.2 方法

肝癌死亡率分布图的建立:在Excel表格中将全省各地3年肝癌死亡资料归总,算出3年平均粗死亡率和标化死亡率,并转换为DBF文件,形成完整的肝癌死亡数据库。

在ArcView 3.0a软件中,以基础地图为背景,导入肝癌死亡数据库,并将数据库中的数据链接到图形属性表中,以肝癌3年平均标化死亡率为参数字段进行分类,链接后,按死亡率大小赋予不同的等级颜色在基础地图上显示出来。

地理资料的GIS空间分析:在ArcView 3.0a软件中将地理数据按地理经度和纬度叠加到基础地图上,应用ArcView 3.0a Spatial Analyst模块功能对各项指标进行空间趋势面分析,并根据指标值的大小进行区域划分,以线条或颜色表示出地理数据的分布。

肝癌及地理数据库的建立:在ArcView 3.0a软件的支持下,以县界地图为基础,从地理空间趋势面分析图中提取出各县的平均地理数据,按地区国际码与肝癌数据库进行匹配联结,将联结好的肝癌数据库和地理数据库以DBF文件形式导出保存。

回归及曲线拟合分析:用SPSS软件对肝癌和各项地理指标作回归分析,进行最佳曲线拟合。

## 2 结果

### 2.1 全省肝癌分布情况

共收集到全省28个县/市1996年至1998年肝癌死亡数据,见表1。经GIS空间分析,形成肝癌死亡分布图<sup>[2]</sup>,从图中可发现苏南和沿海地区如海门、启东等地肝癌死亡率较高,苏中和苏北等地死亡率相对较低。

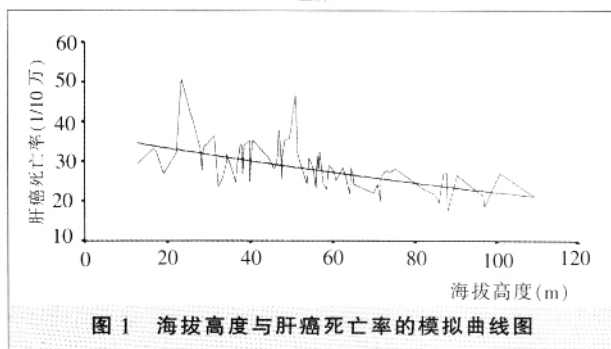
表1 江苏省28个县/市1996年~1998年肝癌标化平均死亡率(1/10万)\*

| 地区 | 标化死亡率 | 地区    | 标化死亡率 | 地区 | 标化死亡率 | 地区  | 标化死亡率 |
|----|-------|-------|-------|----|-------|-----|-------|
| 常熟 | 29.72 | 淮阴楚州区 | 22.94 | 邳州 | 29.33 | 吴县  | 38.06 |
| 常州 | 50.52 | 江阴    | 25.31 | 启东 | 42.61 | 锡山  | 23.25 |
| 大丰 | 31.04 | 金湖    | 35.94 | 苏州 | 55.56 | 新沂  | 24.09 |
| 东海 | 18.51 | 金坛    | 21.94 | 太仓 | 27.60 | 徐州  | 17.90 |
| 赣榆 | 15.96 | 昆山    | 32.55 | 铜山 | 22.30 | 扬中  | 20.16 |
| 海安 | 41.54 | 溧阳    | 25.17 | 无锡 | 18.07 | 宜兴  | 35.87 |
| 海门 | 32.21 | 南京    | 26.66 | 吴江 | 29.83 | 张家港 | 32.17 |

\*:标化平均死亡率以1982年全国人口标化。

### 2.2 肝癌与海拔高度的关系

由图1可见,肝癌死亡率随海拔高度的增加而降低,二者呈显著负相关。回归分析显示:回归系数 $R=0.58$ ,决定系数 $R_2=0.33$ ,校正决定系数 $R_2=0.32$ 。方差分析 $F=37.49$ ,显著性检验 $P<0.0001$ 。回归方程为: $y=36.92e^{-0.0051x}$ ( $y$ 为肝癌死亡率, $x$ 为海拔高度)。



### 2.3 肝癌与温度的关系

肝癌死亡率与6~8月平均温度呈抛物线关系,方差检验有显著相关,而与3~5月平均温度未见有相关(见图2)。回归分析显示:回归系数 $R=0.36$ ,决定系数 $R_2=0.13$ ,校正决定系数 $R_2=0.11$ 。方差检验 $F=5.51$ ,显著性检验 $P=0.006$ 。回归方程: $y=-1757.13+140.63x-2.77x^2$ ( $y$ 为肝癌死亡率, $x$ 为6~8月平均温度)。

### 2.4 肝癌与降雨量及地面蒸发量的关系

肝癌死亡率与年平均降雨量、年平均地面蒸发量、6~8月和9~11月平均降雨量、月地面蒸发量均显著相关,与6~8月降雨量和地面蒸发量的差值呈抛物线关系(见图3),方差检验有显著性。回归系数 $R=0.50$ ,决定系数 $R_2=0.25$ ,校正决定系数 $R_2=0.23$ 。方差分析 $F=12.30$ ,显著性检验 $P<0.0001$ 。

回归方程: $y=-42.75+0.84x-0.002x^2$ ( $y$ 为肝癌死亡率, $x$ 为降雨量与地面蒸发量之差)。

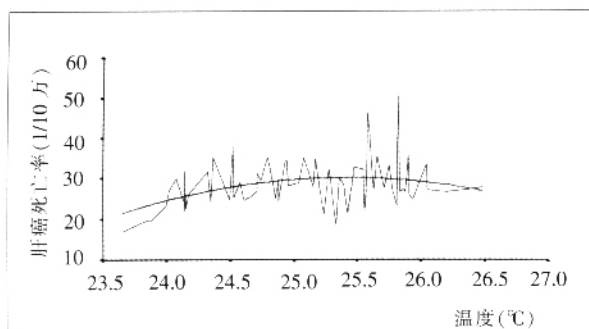


图2 温度与肝癌死亡率的模拟曲线图

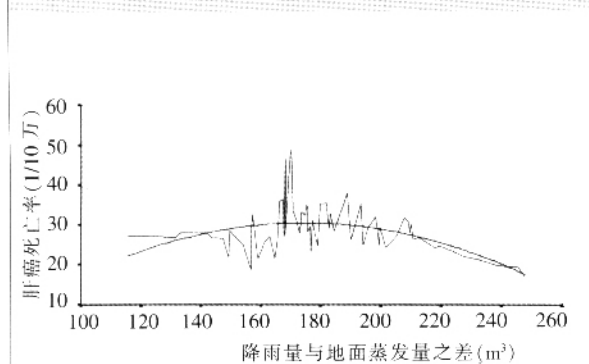


图3 降雨量和地面蒸发量的差值与肝癌死亡率的模拟曲线图

### 3 讨论

地理信息系统是以空间数据库为基础,在计算机软硬件的支持下,对空间相关数据进行采集、管理、操作、分析、模拟和显示,采用地理模型分析方法,适时提供多种空间和动态的地理信息,为研究和决策服务而建立起来的计算机系统。也可简单地定义为用于采集、存储、处理、分析和显示地理空间数据的计算机系统<sup>[3,4]</sup>。此外,GIS还可以充分利用现有多方面的资料进行流行病学分析,从而更好地研究地理、环境与疾病之间的关系。据估计,80%的流行病学研究资料都具有空间属性,如疾病发病情况、温湿度、降雨等都与地理位置有关,GIS可以通过空间关系(地理坐标)将这些数据库连接起来,进行交互显示和分析,同时也为进一步的统计分析提供了基础。本文将肝癌死亡率与地理因素在SPSS软件中做回归分析,拟合最佳曲线,以期得到准确的

回归方程。

人类疾病与健康因不同地区、不同时间而异,越来越多的学者日益关注地理环境与健康、疾病之间的关系。地理环境包括物理性的、生物性的和社会性的地理因素。地球表面的陆地,根据形态特征和海拔高度分为平原、山地、丘陵、高原和盆地五种类型,加上不同气候带的影响,形成了不同的环境特征,对人体健康产生不同影响。研究发现,我省肝癌的地区分布与海拔高度明显相关,平原低地是肝癌的高发区。另外,地貌对水源保护和地下水污染有一定的关系<sup>[5]</sup>。目前国内外有关肝癌与海拔之间关系的研究较少,究其原因可能与海拔影响环境有关。

气温对人体的影响是多方面的,既影响人体生理反应、发病率和死亡率,也与环境污染有密切关系。水气凝结成水滴,降落地面,形成降水。地面水在一定温度作用下,形成蒸发。某地区降水量与地面蒸发量可在一定程度上反应空气中相对湿度的大小。本研究结果显示,肝癌死亡率与温度和降雨量、地面蒸发量都有显著关联,这可能与温度、降雨量(或相对湿度)影响黄曲霉生长有关。作为肝癌病因假说之一的黄曲霉毒素,需要在一定的季节、一定的温湿度或降雨量,尤其是6~8月的气候条件下才能很好地生长繁殖。

然而,肝癌的发生和发展是一个多因素、多步骤的过程,本文仅对几项地理因素与肝癌死亡率的关系作了简单的单因素分析,有待在以后的研究中作进一步的多因素分析。

### 参考文献:

- [1] 卜绍唐,熊宗璠.江苏省人口全死因研究(1990~1992年)[M].南京:东南大学出版社,1996.42~47.
- [2] 胡晓抒,周晓农,孙宁生,等.江苏省恶性肿瘤分布态势地理信息系统的空间分析[J].中华流行病学,2002,23(1):73~74.
- [3] 韩光红,张习坦.地理信息系统及其在流行病学研究中的应用[J].中华流行病学杂志,2001,22(1):65~67.
- [4] 孙宁生,胡晓抒,周晓农.地理信息系统与卫生管理[J].江苏卫生事业管理,1999,10(48):59~62.
- [5] 耿贯一.流行病学[M].第二版.北京:人民卫生出版社,1996.352~354.