

影响乳腺癌发生的空间危险因素分析

彭小冬¹,彭琳²,吴库生¹

(1.汕头大学医学院预防医学教研室,广东 汕头 515041;
2.汕头大学医学院附属肿瘤医院,广东 汕头 515031)

摘要:[目的]在地理信息系统支持下探讨乳腺癌死亡率与空间环境因素之间的关系。**[方法]**在Excel中建立1/10抽样地区乳腺癌死亡率数据库(1990~1992年),在Arc/Info 9.0中连接已制作的抽样地区电子地图。利用空间分析的方法提取抽样地区分县的相应空间环境数据(1961~1990年/月平均数据)。对乳腺癌死亡率与相应地区空间环境因素进行Pearson或Spearman相关分析;对环境数据进行因子/主成分分析后与死亡率数据进行多元逐步回归分析。**[结果]**空间分布地图显示中国乳腺癌死亡率分布呈现一定的分布规律,具有东北部向中部、西南部递减的趋势。相关分析显示乳腺癌标准化死亡率与1~12月月平均风速呈正相关(相关系数按月份顺序分别为0.164,0.137,0.135,0.141,0.153,0.136,0.134,0.143,0.193,0.186,0.183,0.156,P<0.05),与海拔呈负相关($r_s=-0.159$,P<0.05)。经因子/主成分分析后多元逐步回归建模,最终1~12月月平均风速进入模型。**[结论]**中国乳腺癌在空间分布上有明显的趋势,乳腺癌死亡率与风速、海拔相关。

关键词:地理信息系统;乳腺肿瘤;气候;地理

中图分类号:R737.9 文献标识码:A 文章编号:1004-0242(2015)10-0805-06
doi:10.11735/j.issn.1004-0242.2015.10.A001

An Analysis of Spatial Risk Factors on Breast Cancer by Using Geographic Information System

PENG Xiao-dong¹, PENG Lin², WU Ku-sheng¹

(1. Department of Preventive Medicine, Shantou University Medical College, Shantou 515041, China;
2. The Cancer Hospital of Medical College of Shantou University, Shantou 515031, China)

Abstract: [Purpose] To explore the relationship between breast cancer mortality and ambient climate, geographic factors using geographic information system (GIS). [Methods] Database of breast cancer mortality of 1/10 sampling areas surveyed in 1990~1992 was established in Excel and linked with the digital polygon maps of study areas in Arc/Info 9.0. Geographic and climate data of sampling areas were extracted from the raster dataset (1961~1990, annual/monthly average data). Pearson correlation analysis or Spearman correlation analysis and multiple regression analysis after factor analysis were carried out to analyze the relationship between breast cancer mortality and the spatial environment factors. [Results] Spatial distribution maps showed that China breast cancer mortality distribution had regularities of distribution, which decreased from northeast, middle to southwest. Correlation analysis showed that breast cancer standardization mortality was positively correlated with average wind speed from January to December (According to the month order, correlation coefficient was 0.164, 0.137, 0.135, 0.141, 0.153, 0.136, 0.134, 0.143, 0.193, 0.186, 0.183, 0.156, respectively. $P<0.05$), but negatively correlated with altitude ($r_s=-0.159$, $P<0.05$). After factor/principal component analysis and multiple stepwise regression analysis, the variable average wind speed from January to December finally entered the regression model. [Conclusions] Breast cancer had significant spatial aggregation in China, and are correlated with average wind speed and altitude.

Key words: geographic information system; breast cancer; climate; geography

乳腺癌是女性常见的恶性肿瘤之一,严重威胁

收稿日期:2015-03-06;修回日期:2015-06-03

基金项目:汕头大学医学院临床科研提升计划(201413);
国家自然科学基金(81470152)

通讯作者:吴库生,E-mail:kswu@stu.edu.cn

着女性健康,现已成为发展中国家女性死亡的首位恶性肿瘤^[1]。据资料统计,我国2008年女性乳腺癌发病16.9万例,是女性第二位常见的恶性肿瘤;2008年我国女性乳腺癌死亡约4.5万例,是女性第

六位恶性肿瘤死亡原因^[2]。中国女性乳腺癌发病率和死亡率在全球处于较低的水平,但呈迅速增长的趋势,尤其是农村地区近10年来上升趋势明显^[3]。

乳腺癌的地理分布提示其与环境因素有关,有人提出环境因素占癌症病因的80%~90%^[4]。土壤、植被、地形和气候因素是人类赖以生存的基本生态环境,它们之间息息相关。对于降水量、蒸发量、气温等因素与肝癌的关系^[5],以及土壤、植被等地理因素与食管癌的关系^[6,7]国内早有研究。

随着地理信息系统(geographic information system, GIS)技术的出现与发展,空间数据挖掘技术和方法的不断进步,使对空间数据的处理与分析变得相对成熟,GIS在卫生领域尤其在流行病学研究领域中得到了越来越多的应用^[8]。

基于GIS的强大空间数据处理与分析能力、交互能力以及图形表达能力,本研究将对一些气候、地理空间因素进行提取分析,并结合相应地区乳腺癌死亡率数据,分析影响乳腺癌发生、死亡与气候、地理空间因素相关性。

1 资料与方法

1.1 疾病数据

乳腺癌死亡率数据来自于全国肿瘤防治办公室1990~1992年全国1/10人口抽样地区死因回顾调查。数据包括乳腺癌粗死亡率、1964年中国人口标化率、1976年世界人口标化率。由于全国行政区划进行了多次调整,因此按2013年行政区划调整校对后,在Excel中建立抽样地区乳腺癌死亡率数据库,并输入ArcGIS 9.0(ESRI, Inc., Redlands, CA)软件以建立GIS数据库备用。

1.2 电子地图制作及与疾病数据联接

全国基本电子地图(1:4 000 000)由国家测绘局提供,为ESRI格式的shape文件。以此为基础地图在Arc/Info 9.0中加载后通过在属性表中选择行政代码为抽样研究的地区,形成抽样地区县界多边形电子地图。

用此抽样地区县界多边形电子地图与乳腺癌死亡率数据库连接,然后再与相关栅格数据进行空间分析,得到各抽样地区的相关空间因素的平均数据。

1.3 气候及地理数据

原始数据:中国地形图,1961~1990累年平均降水量、蒸发量,1971~2000年累年平均降水量,2000年平均降水量,1961~1990年累年平均气温,1961~1990年累年平均最高/最低气温,1951~2000年累年各月平均风速等栅格数据来自于中国自然资源数据库。

空间分析分县数据:利用抽样地区县界多边形电子地图在ArcGIS 9.0的空间分析模块中通过区域统计的方法(以县界地图为分区图层,以上述各栅格数据为值图层)提取抽样地区各市县的上述地理环境数据平均值,并在此基础上计算干燥指数(计算公式为年平均蒸发量/年平均降水量)与水热指数(计算公式为年平均降水量/年平均气温)。

海拔高程数据:各抽样点的海拔数据由Google Earth中文版软件(6.1.0.5001版)查询得到。利用抽样点的县或区名称在Google Earth中查询,软件会定位到相应地区,利用鼠标对准相应地区的政府所在地,软件会显示相应的经度、纬度及海拔数据。

1.4 统计学处理

1.4.1 相关分析

由于部分空间变量(1月平均蒸发量、2月平均蒸发量、3月平均蒸发量、12月平均蒸发量、海拔)为非正态分布变量,因此采用Spearman等级相关分析方法,其余空间变量相关性分析采用Pearson相关分析,计算中国女性乳腺癌标化死亡率与空间因素的相关或等级相关系数,并进行显著性检验。

1.4.2 因子分析、主成分分析与多元回归分析

为消除变量间的共线性问题,对提取的空间因素首先进行因子/主成分分析。然后以乳腺癌标化死亡率的对数为因变量,以因子分析后析出的因子为自变量,进行多元逐步回归分析(纳入水准0.05,排除水准0.10),以分析各空间因素与乳腺癌死亡率之间的综合关系。以上分析采用SPSS17.0(SPSS Inc., Chicago, Illinois)统计包。

2 结 果

2.1 中国乳腺癌死亡率空间分布

在Arc/Info 9.0中通过反距离权重插值法(inverse distance weighting, IDW)来描述中国女性乳腺

癌的空间分布,从Figure 1中可看出中国女性乳腺癌死亡率在地域上呈现一定的分布规律,东北部向中部、西南部递减的趋势,其中乳腺癌标化死亡率最高的有绥化县、厦门市鼓浪屿区、湛江市赤坎区、平顺县、青岛市市北区,最低的有乌审旗、南城县、崇义县、桓台县、仙居县。

2.2 地理气候因素与乳腺癌死亡率的相关分析

对乳腺癌标化死亡率与环境空间因素之间的相关关系进行Pearson或Spearman相关分析,结果显示1~12月平均风速与女性乳腺癌死亡率呈正相关(相关系数按月份顺序分别为0.164,0.137,0.135,0.141,0.153,0.136,0.134,0.143,0.193,0.186,0.183,0.156,除9月、10月、11月为 $P<0.01$ 外,其余均 $P<0.05$),与平均海拔呈负相关($r_s=-0.159$, $P<0.05$),其他相关无统计学意义(Table 1)。

2.3 因子/主成分和多元逐步回归分析

对提取的环境空间因素进行KMO检验及球形检验显示,KMO统计量为0.702,Bartlett球形检验

拒绝单位相关阵的原假设($P<0.001$),提示适合做因子分析。因子数目与特征值散点图如Figure 2所示。6个因子/主成分(由于加入年均日照和年均蒸发量后显示不是正定矩阵,故去掉这两个变量)被描述出,其中因子1代表(1,2,3,4,5,8,9,10,11,12月、年总)蒸发量、降水量、气温;因子2代表日照时间、6,7月蒸发量、干燥指数;因子3代表1~12月月平均风速;因子4代表海拔、1974年月平均气温分级;因子5代表水热指数;因子6代表其他因素。以此6个因子/主成分与乳腺癌标化死亡率对数进行多元逐步回归分析后,结果因子3进入方程。其决定系数、方差分析及参数估计见Table 2。由Table 2可见,进入方程的最终变量为1~12月月平均风速。

3 讨 论

在西方发达国家,乳腺癌发病率处于较高水平,乳腺癌死亡率也位居女性癌症死亡的首位,而包括

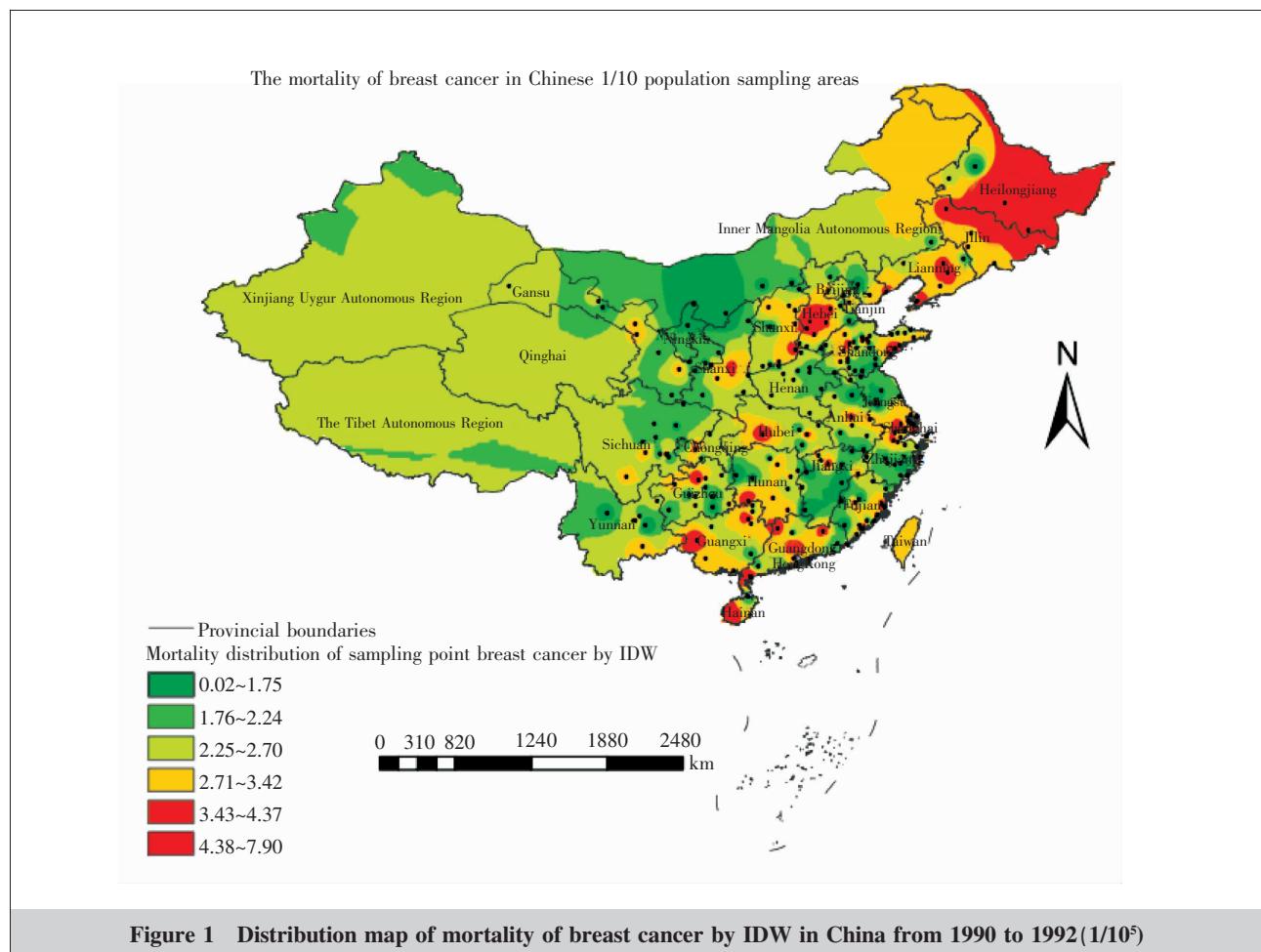


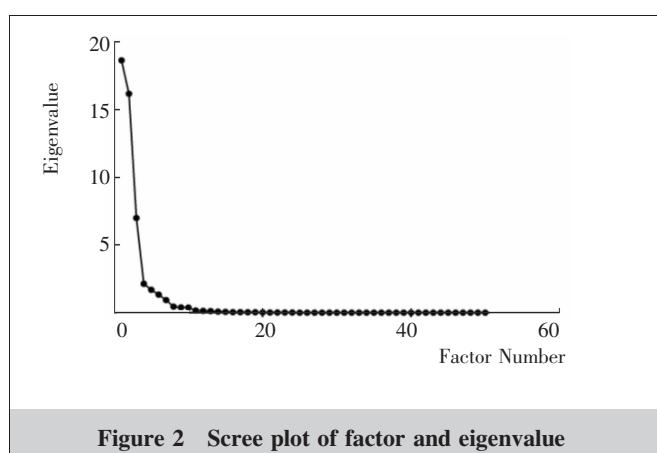
Table 1 Correlation analysis between mortality of breast cancer and the spatial environment factors in China

The spatial environment factors of sampling point	Standardization mortality of breast cancer		The spatial environment factors of sampling point	Standardization mortality of breast cancer	
	r	P		r	P
Year average rainfall during 30 years(1961~1990)(mm)	0.009	0.893	Average wind velocity in March*	0.135	0.036
Year average rainfall during 30 years(1971~2000)(mm)	0.008	0.907	Average wind velocity in April*	0.141	0.028
Year average rainfall in 2000 (mm)	0.000	0.998	Average wind velocity in May*	0.153	0.017
Year average temperature of 30 years (°C)	0.022	0.728	Average wind velocity in June*	0.136	0.035
Year average lowest temperature of 30 years (°C)	-0.010	0.880	Average wind velocity in July*	0.134	0.037
Year average highest temperature of 30 years (°C)	-0.046	0.476	Average wind velocity in August*	0.143	0.026
Year average temperature in 1978 (°C)	-0.062	0.335	Average wind velocity in September**	0.193	0.003
Average altitude (m) [△] *	-0.159	0.013	Average wind velocity in October**	0.186	0.004
Month average temperature grading during 30 years	-0.019	0.767	Average wind velocity in November**	0.183	0.004
Month average temperature grading in August 1974	0.055	0.392	Average wind velocity in December*	0.156	0.015
Total evaporation capacity of year (mm)	-0.018	0.780	Total sunshine hours of year(h)	0.053	0.415
Year average evaporation capacity (mm)	0.062	0.340	Year average sunshine hours(h)	0.090	0.163
Average evaporation capacity in January [△]	-0.011	0.863	Average sunshine hours in January	0.075	0.244
Average evaporation capacity in February [△]	-0.011	0.861	Average sunshine hours in February	0.077	0.232
Average evaporation capacity in March [△]	-0.022	0.736	Average sunshine hours in March	0.081	0.209
Average evaporation capacity in April	0.015	0.817	Average sunshine hours in April	0.073	0.255
Average evaporation capacity in May	-0.096	0.138	Average sunshine hours in May	0.089	0.17
Average evaporation capacity in June	0.079	0.221	Average sunshine hours in June	0.062	0.337
Average evaporation capacity in July	0.053	0.408	Average sunshine hours in July	0.036	0.576
Average evaporation capacity in August	0.040	0.537	Average sunshine hours in August	0.032	0.622
Average evaporation capacity in September	0.024	0.715	Average sunshine hours in September	0.118	0.067
Average evaporation capacity in October	0.058	0.367	Average sunshine hours in October	0.098	0.129
Average evaporation capacity in November	0.064	0.324	Average sunshine hours in November	0.090	0.162
Average evaporation capacity in December [△]	-0.008	0.898	Average sunshine hours in December	0.077	0.234
Average wind velocity in January*	0.164	0.011	Drought index	0.001	0.988
Average wind velocity in February*	0.137	0.033	Water heat index (mm/°C)	0.087	0.177

Annotation: [△]:spearman correlation coefficient; *:P<0.05; **:P<0.01; Average evaporation capacity in Jan. ~ Dec. (mm) was calculated based on the mean data of 30 years from 1971 to 2000; Average wind velocity in Jan. ~ Dec. was calculated based on the mean data of 50 years; Average sunshine hours in Jan. ~ Dec. (h) was calculated based on the mean data of 25 years.

Table 2 Multiple stepwise regression analysis between breast cancer mortality and the spatial environment factors in China

Environment factors	Unstandardized Coefficients	Std. Error	Standardized Coefficients	R ²	Adjusted R ²	F	P
Factor 3	0.033	0.014	0.152	0.023	0.019	5.637	0.018

**Figure 2 Scree plot of factor and eigenvalue**

中国在内的大多数亚洲国家乳腺癌发病率和死亡率相对较低一些^[9,10]。这种地域间的差异提示人种(基因)和环境(包括生活习惯)可能影响乳腺癌的发生^[11]。对乳腺癌的地理分布特征及影响因素进行分析,在乳腺癌病因学及其防治研究中具有重要作用。通过绘制疾病空间分布地图,可了解疾病空间分布特征与规律,为病因学研究提供依据。GIS为绘制交互式的疾病分布地图提供了有用的工具,不仅可直观地显示疾病分布特征,而且可进行交互式查询、与环境因素间的叠加分析等^[12]。

本次研究属于生态学研究,运用 GIS 分析疾病分布,为疾病预防控制提供一种新的病因学分析方法。经分析发现我国女性乳腺癌死亡率在地域上呈现一定的分布规律,东北部向中部、西南部递减的趋势,与相关报道的结果一致。娄清涛等^[13]的研究提出乳腺癌死亡率具有明显的地理空间分布特征;有报道,在地理分布上,乳腺癌死亡率呈现由西南向东北、西部向东部增加的趋势^[14]。亦有研究^[15]认为是沿东部、中部和西部顺序依次降低的态势。但由于国内环境空间数据的获取与处理技术的难易程度和疾病登记制度的不完善,目前对环境空间因素与乳腺癌的研究较少见,本研究在参考前人研究^[16]的基础上,用环境空间因素与乳腺癌死亡率资料在 GIS 技术的支持下进行分析。

乳腺癌发生与死亡的流行因素是遗传因素、生活方式和环境暴露等多种因素及其相互作用的结果^[2]。气候因素为人类生存的基本生态环境之一,因此研究气候等地理环境因素对于人类疾病的影响具有重要的意义。气候因素持续与其他生态及社会因素相交互作用扩大了某些疾病的流行^[17]。本次研究表明风速与乳腺癌死亡率之间存在正相关性,经因子/主成分分析后多元逐步回归,最终进入模型的变量是 1~12 月月平均风速。在日常生活中,风速和风向,结合当地的地形、气候特征、工业化生产情况可能对其他环境因素如降水量、蒸发量、植被覆盖、水土流失、生物有害物质等起到间接作用,对人体健康具有很大的潜在影响,如季风、沙尘暴、寒潮;且风速过大会加快污染物向周围扩散的速度,如大气中颗粒物浓度的改变,SO₂、NO_x 等物质向空中扩散后沉降下来,增加对人体暴露的机会,从而造成健康危害^[18]。亦有报道风速的大小和方向在一定程度上影响环境致癌污染物在空中的扩散与传播,如挥发性有机化合物^[19]、多环芳烃^[20],还有放射性物质^[21]等。但环境空间因素影响乳腺癌发生与死亡的具体机制仍未知,应进一步深入探讨验证。

另外,本研究显示海拔高度与女性乳腺癌死亡率之间呈负相关。目前关于海拔高度与癌症的关系国内外有较多的报道,对海拔高度引起癌症的说法还存在争议。如 Hart 等^[22]显示高海拔地区癌症死亡率低,而低海拔地区癌症死亡率高,并认为有辐射兴奋效应的存在。Daniela 等^[23]的研究认为高海拔地

区紫外线暴露增加从而导致皮肤黑色素瘤的发生率增加,并且在一定程度上与气候改变有关。Krain^[24]等的研究也认为高海拔暴露及飞行驾驶与皮肤、睾丸、膀胱、甲状腺等肿瘤有密切关系。而在本次研究中,经因子/主成分分析后进行多元逐步回归分析,海拔因素并没有进入模型,可能原因是海拔可能影响其他的地理环境如地质、气候等,从而对乳腺癌死亡率产生影响。

全国行政区划不断调整,部分抽样地区市区县已改变,致使抽样地区资料无法全部得到,这可能是出现缺失的原因之一。另外,本次研究未分析土壤构成、植被覆盖、空气与水污染、太阳净辐射资料、烟草消耗和 GDP 等相关数据,不能全面地分析环境空间因素对乳腺癌死亡率的影响。相信对未来空间数据库的建立进行全面分析,也可结合 GIS 与遥感技术共同分析,会有更好的结果。

中国乳腺癌的分布具有明显的地域性,提示地理环境因素在乳腺癌的发生发展过程中起到一定的作用。本研究初步揭示了地理环境因素对于乳腺癌死亡率的影响,为其他有关乳腺癌的病因研究提供参考依据。

参考文献:

- [1] Jemal A,Bray F,Center M,et al. Global cancer statistics [J]. CA Cancer J Clin,2011,61(2):69–90.
- [2] Zhang ML,Huang ZZ,Zheng Y. Estimates and prediction on incidence,mortality and prevalence of breast cancer in China,2008 [J]. Chin J Epidemiol,2012,33(10):1049–1051.
[张敏璐,黄哲宙,郑莹.中国 2008 年女性乳腺癌发病、死亡和患病情况的估计及预测[J].中华流行病学杂志,2012,33(10):1049–1051.]
- [3] Zheng Y,Wu CX,Zhang ML. The epidemic and characteristics of female breast cancer in China[J]. China Oncology,2013,23(8):561–569.[郑莹,吴春晓,张敏璐.乳腺癌在中国的流行状况和疾病特征 [J]. 中国癌症杂志,2013,23(8):561–569.]
- [4] Wu SC. Environment and Health [M].Beijing:People's Medical Publishing House,1982.230.[吴沈春.环境与健康[M].北京:人民卫生出版社,1982.230.]
- [5] Qin Y,Hu XS,Zhao JK,et al. Regression analysis on liver cancer and geographic factors in Jiangsu province[J]. Bulletin of Chinese Cancer,2003,12(11):10–12.[覃玉,胡晓抒,赵金扣,等.江苏省肝癌与地理因素回归分析[J].中

- 国肿瘤,2003,12(11):10-12.]
- [6] Wu KS,Huo X,Zhu GH. Soil and vegetation characteristics from high- and low-risk areas of esophageal cancer in China[J]. Asian Pac J Cancer P,2007,8(4):597-601.
- [7] Wu KS,Li K. Association between esophageal cancer and drought in China by using Geographic Information System [J]. Environ Int,2007,33(5):603-608.
- [8] Zhang RJ,Tao HX. Research on mining method of spatial data based on FNN[J]. Journal of Shandong University of Science and Technology(Nature Science),2003,22(4):24-27.[张瑞菊,陶华学. 基于模糊神经网络的空间数据挖掘方法研究 [J]. 山东科技大学学报 (自然科学版), 2003, 22(4):24-27.]
- [9] Qin LQ,Xu JY,Wang PY,et al. Effects of milk and its products on breast cancer risk:a review [J]. Chin J Cancer Prev Treat,2007,14(17):1345-1349.[秦立强,徐加英,王培玉,等. 牛奶及其制品与乳腺癌关系的研究进展[J]. 中华肿瘤防治杂志,2007,14(17):1345-1349.]
- [10] Liu FW,Liu SZ,Quan PL,et al. An analysis of mortality with female breast cancer from 1984 to 2002 in Henan province[J]. Bulletin of Chinese Cancer,2004,13(5):27-29.[刘法文,刘曙正,全培良,等. 1984~2002年河南省女性乳腺癌死亡率分析[J]. 中国肿瘤,2004,13(5):27-29.]
- [11] Qi YJ. Analysis on high risk factors of the cause of breast cancer[J]. Qinghai Medical Journal,2009,39(4):71-74.[祁玉娟. 乳腺癌病因的高危因素分析 [J]. 青海医药杂志, 2009, 39(4):71-74.]
- [12] Wu KS,Li K. Study on geo-epidemiology of esophageal cancer in guangdong province by using Geographic Information System [J]. Journal of Shantou University Medical College,2006,19(4):230-232.[吴库生,李克. 基于地理信息系统的广东省食管癌地理流行病学研究[J]. 汕头大学医学院学报,2006,19(4):230-232.]
- [13] Lou QT,Zhang XF,Quan PL,et al. Trend-surface analysis on the geographical distribution of mortality with breast cancer in Henan province [J]. China Cancer,2010,19(9):569-572.[娄清涛,张欣峰,全培良,等. 河南省乳腺癌死亡率地理分布的趋势面分析[J]. 中国肿瘤,2010,19(9):569-572.]
- [14] Shi XJ,Zhang XJ,Wang FS,et al. Analysis on mortality distribution of female breast cancer in China,1991- 2010 [J]. Chin J Dis Control Prev,2012,16 (9):743-747.[石晓君,张晓佳,王富生,等. 1991-2010年中国女性乳腺癌的死亡分布特征 [J]. 中华疾病控制杂志,2012,16(9):743-747.]
- [15] Zheng Y,Wu CX,Wu F. Status and trends of breast cancer mortality in Chinese female [J]. Chin J Prev Med, 2011,45(2):150-154.[郑莹,吴春晓,吴凡. 中国女性乳腺癌死亡现况和发展趋势 [J]. 中华预防医学杂志, 2011,45(2):150-154.]
- [16] Wu KS,Li K,Yu P. Application of GIS on epidemiology of cancer [J]. Journal of Shantou University Medical College,2006,19(1):55-57.[吴库生,李克,于萍. 地理信息系统在肿瘤流行病学中的应用 [J]. 汕头大学医学院学报, 2006, 19(1):55-57.]
- [17] Acuna-Soto R,Stahle DW,Cleaveland MK,et al. Megadrought and megadeath in 16th century Mexico [J]. Emerg Infect Dis,2002,8(4):360-362.
- [18] Bernard SM,Samet JM,Grambsch A,et al. The potential impacts of climate variability and change on air pollution-related health effects in the United States [J]. Environ Health Persp,2001,109(Suppl2):199-209.
- [19] Kountouriotis A,Aleiferis PG,Charalambides AG. Numerical investigation of VOC levels in the area of petrol stations [J]. Sci Total Environ,2014,470-471:1205-1224.
- [20] Bandowe BA,Meusel H,Huang R,et al. PM2.5-bound oxygenated PAHs,nitro-PAHs and parent-PAHs from the atmosphere of a Chinese megacity:seasonal variation, sources and cancer risk assessment[J]. Sci Total Environ, 2014,473-474:77-87.
- [21] Drozdovitch V,Zhukova O,Germenckuk M,et al. Database of meteorological and radiation measurements made in Belarus during the first three months following the Chernobyl accident[J]. J Environ Radioactiv, 2013 , 116:84-92.
- [22] Hart J. Cancer mortality for a single race in low versus high elevation counties in the U.S. [J]. Dose Response, 2011,9(3):348-355.
- [23] Haluza D,Simic S,Moshammer H. Temporal and spatial melanoma trends in Austria:an ecological study [J]. Inter J Env Res Pub Heal,2014,11(1):734-748.
- [24] Krain LS. Aviation,high altitude,cumulative radiation exposure and their associations with cancer [J]. Med Hypotheses,1991,34(1):33-40.