

## 地理环境对健康人血清脂联素水平的影响\*

纪旭<sup>1</sup>, 葛森<sup>1</sup>, 李小平<sup>1</sup>, 井静<sup>1,2</sup>, 林宁<sup>1</sup>

(1.陕西师范大学 地理科学与旅游学院,陕西 西安 710119;

2.宝鸡文理学院 地理与环境学院地理系,陕西 宝鸡 721016)

**摘要:** 分析我国健康人血清脂联素水平在空间上的分布规律特征,探究地理环境对其影响作用,并为制定不同地区血清脂联素的标准提供科学的依据.通过搜集到的全国 106 个市县级单位 13 988 例健康人血清脂联素参考值,运用空间自相关、相关分析和地理探测器方法研究其与地理环境指标之间的相关性.对比岭回归和支持向量机预测模型后,选择最优模型进行建模.运用趋势分析,对 APN 参考值的预测数据进行空间分析.结果表明,我国健康人血清脂联素水平与纬度、海拔高度、年平均气温、年降水量、表层土壤 pH 和表层土壤总可交换量 6 项地理指标存在相关性,最优预测模型回归方程为 $Y=11.9905-0.0134X_2-0.00092X_3+0.0735X_5+0.00059X_6-0.10201X_{10}-0.06712X_{11} \pm 4.218$ .我国健康人血清脂联素水平整体上呈东南高西北低,沿海高于内陆的特征.

**关键词:** 血清脂联素;地理环境;地理探测器;岭回归分析;支持向量机

**中图分类号:** R188 **文献标志码:** A **文章编号:** 1007-9793(2019)02-0060-06

脂联素(adiponectin, APN)是脂肪细胞分泌的一种内源性生物活性多肽,是脂肪细胞表达最丰富的蛋白质产物之一.人体的研究发现,APN 是一种胰岛素增敏激素,也是一种抗炎因子,具有抗动脉粥样硬化、抗糖尿病以及抗炎反应等特性. APN 水平的降低预示着高胰岛素血症和高血糖症的发生,提示 II 型糖尿病和冠心病的发展.临床研究发现<sup>[1-4]</sup>,正常人血液循环中有丰富的脂联素,其浓度约为 1.9~17.0 mg/L,约占血浆蛋白总量的 0.01%. APN 在人体中的浓度由基因和环境共同决定,而在制定医学参考值标准时,却忽略了地理环境因素对参考值的影响.现已有关于民族、性别、年龄和体重等对 APN 水平影响的研究<sup>[5-6]</sup>,但地理环境对其影响的研究较少,因此研究地理环境对 APN 水平的影响,探究其影响机制,对于完善医学参考值标准的制定,增加临床判断准确性方面具有重要的科学价值.

### 1 资料与方法

#### 1.1 资料

##### 1.1.1 APN 参考值的获取

通过检索中国期刊网、万方数据库、维普数据库、中国生物医学文献数据库、超星数字图书馆及中国重要会议数据库等,搜集了 2010~2018 年上半年以来,全国 106 个市、县、区级医院和相关医疗单位测定的 13 988 例健康人 APN 参考值,其中男性 6 734 例(48.1%),女性 7 254 例(51.9%),年龄范围 20~78 岁.根据相关研究,性别、年龄及种族对参考值存在一定影响,但绝大部分研究中仅取平均值,故研究时不区分性别、年龄及种族.资料来源于我国 28 个省、直辖市、自治州,缺少港澳台地区资料.所有实测数据均来自正常人群,纳入标准:  $18.5 \text{ kg/m}^2 \leq \text{BMI 指数} < 24 \text{ kg/m}^2$ ,排除标准:高

\* 收稿日期:2019-01-05

基金项目:国家自然科学基金资助项目(40141002).

作者简介:纪旭(1994-),女,江苏南京人,硕士研究生,主要从事健康地理方面研究.

通信作者:葛森.E-mail:gemiao@snnu.edu.cn.

血压、糖尿病、高脂血症等疾病。所有受试者禁饮食 8~10 h 后于次日清晨空腹采取肘静脉血,利用酶联免疫吸附试验(ELISA)法测定。

### 1.1.2 地理指标的选取

本文选取 4 类地理环境指标,其中空间位置指标:经度  $X_1(^{\circ})$ 和纬度  $X_2(^{\circ})$ ,地势指标:海拔高度  $X_3(\text{m})$ 均来源于国家测绘局数据中心;气象指标:年日照时数  $X_4(\text{h})$ 、年平均气温  $X_5(^{\circ}\text{C})$ 、年降水量  $X_6(\text{mm})$ 、气温年较差  $X_7(^{\circ}\text{C})$ 和年平均相对湿度  $X_8(\%)$ 来自中国气象科学数据共享服务网(<http://cdc.cma.gov.cn/>);土壤指标:表层土壤有机质含量  $X_9(\% \text{wt.})$ 、表层土壤 pH  $X_{10}$ 、表层土壤总可交换量  $X_{11}(\text{cmol/kg})$ 和表层土壤盐分  $X_{12}(\text{g/kg})$ 来源于世界和谐土壤数据库(<http://www.fao.org/nr/land/soils/harmonized-world-soil-database/zh/>)。

## 1.2 方法

### 1.2.1 空间自相关

根据文献<sup>[7-8]</sup>,对数据进行空间自相关分析,依据 Moran's  $I$  统计及相应的  $P$  值和  $Z$  得分来判断数据在空间上是否存在相关性。

### 1.2.2 相关分析

运用相关分析<sup>[9]</sup>,计算出医学参考值与各项地理指标之间的相关系数  $r$  和  $P$  值,筛选出与之相关的地理指标。

### 1.2.3 地理探测器

地理探测器是一种新的统计学方法<sup>[10-12]</sup>,根据地理探测器计算出的  $q$  统计量可用于探测各项地理指标对医学参考值空间分布的解释力大小。

### 1.2.4 共线性诊断

由于共线性问题的广泛存在<sup>[13]</sup>,诊断数据共线性情况可以据此选择适当的方法进行建模,有利于提高预测模型的精确度。

### 1.2.5 岭回归分析

运用岭回归分析<sup>[14]</sup>,对医学参考值和地理指标进行建模分析,可有效解决共线性问题,使模型更符合实际。

### 1.2.6 支持向量机

支持向量机是以统计学理论为基础的新一代机器学习算法<sup>[15-16]</sup>,运用该方法对数据进行建模分

析,能够较好地解决小样本非线性问题。

### 1.2.7 趋势分析

趋势分析是地统计学中一种空间数据分析方法<sup>[17-19]</sup>,将预测的每个点都投影到一个东西向和南北向的正交平面上,来反映特定方向上的存在趋势,根据趋势图反映东西方向上或南北方向上的趋势变化。

## 2 结果

### 2.1 空间自相关结果

利用 ArcGIS 软件进行 Moran's  $I$  指数分析,如图 1 所示,得到全局 Moran's  $I \neq 0, P < 0.05$ ,标准化统计量  $Z$  值为  $-2.058$ ,拒绝零假设,则数据在空间上表现出负相关性,呈离散性分布,即随着空间位置的离散,相关性显著。说明在整个中国区域内,APN 参考值在空间上存在自相关性。

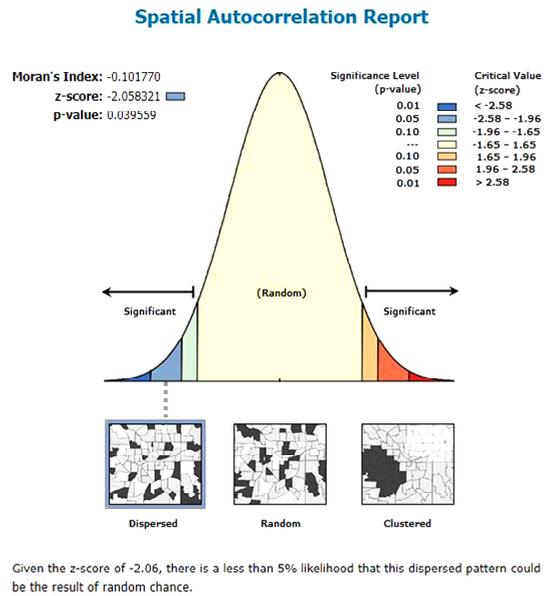


图 1 APN 的空间自相关分析

Fig.1 Spatial autocorrelation analysis of APN

### 2.2 相关分析

运用 SPSS19.0 中的 Spearman 相关分析计算出 APN 参考值与 12 项地理指标的相关系数  $r$  和  $P$  值,结果见表 1.共有 6 项地理指标与 APN 参考值有相关性,分别为  $X_2, X_3, X_5, X_6, X_{10}, X_{11}$ .其中与  $X_2, X_3, X_{10}, X_{11}$  呈负相关关系,与  $X_5, X_6$  呈正

相关关系.

表 1 血清脂联素参考值与地理指标的相关系数

Table 1 Correlation coefficient between APN and the geographical factors

地理指标	<i>r</i>	<i>P</i>
X <sub>1</sub>	0.062	0.516
X <sub>2</sub>	-0.199*	0.034
X <sub>3</sub>	-0.244**	0.009
X <sub>4</sub>	-0.161	0.087
X <sub>5</sub>	0.228*	0.015
X <sub>6</sub>	0.229*	0.014
X <sub>7</sub>	-0.160	0.090
X <sub>8</sub>	0.161	0.087
X <sub>9</sub>	-0.044	0.645
X <sub>10</sub>	-0.185*	0.049
X <sub>11</sub>	-0.223*	0.017
X <sub>12</sub>	-0.021	0.828

\*\*表示相关性极显著,  $P < 0.01$ ; \*表示相关性显著,  $P < 0.05$

### 2.3 地理探测器

以 APN 参考值为因变量, 与之相关的地理指标为自变量, 运用地理探测器进行分析, 确定各地理指标对 APN 参考值空间分布的解释力大小. 采用四分法对与之相关的 6 项地理指标数据进行离散化处理后, 根据因子探测的计算结果  $q$  值大小, 将 6 项地理指标对 APN 参考值的空间分布影响力由强到弱排序:  $X_3 (q=0.0827) > X_6 (q=0.0755) > X_{11} (q=0.0502) > X_{10} (q=0.0382) > X_5 (q=0.0377) > X_2 (q=0.0362)$ , 因此认为  $X_3$  为影响 APN 参考值空间分布的主要因子.

交互探测的结果显示:  $X_2 \cap X_5$ 、 $X_2 \cap X_6$ 、 $X_5 \cap X_6$  的交互探测的结果为双因子增强, 其余指标交互作用后的结果均为非线性增强, 说明任意两种地理指标的交互作用对 APN 参考值的空间分布都要强于单一地理指标的独立作用.

### 2.4 共线性诊断结果

运用 SPSS19.0 对地理指标进行共线性诊断, 结果见表 2.  $X_2$  和  $X_5$  的方差膨胀因子  $> 10$ , 容忍

度接近于 0, 提示存在严重的共线性,  $X_6$ 、 $X_{10}$  和  $X_{11}$  的条件指数  $> 10$ , 特征值接近于 0, 提示存在多重共线性问题.

表 2 APN 参考值共线性诊断结果

Table 2 The multicollinearity result of APN

地理指标	方差膨胀因子	容忍度	特征值	条件指数
X <sub>2</sub>	28.069	0.036	0.842	2.626
X <sub>3</sub>	2.312	0.433	0.211	5.240
X <sub>5</sub>	20.820	0.048	0.114	7.128
X <sub>6</sub>	6.719	0.149	0.021	16.540
X <sub>10</sub>	3.936	0.254	0.004	36.210
X <sub>11</sub>	3.572	0.280	0.000	137.163

### 2.5 岭回归分析

运用 SAS 软件, 通过编程进行岭回归分析, 得到 APN 参考值岭迹图, 如图 2. 图 2 中不同的曲线代表不同的自变量, 由岭迹图可知, 当  $K=0.4$  时, 岭迹曲线逐渐趋于稳定, 此时各系数的方差膨胀因子都在 10 以内, 相比于最小二乘法, 此时的均方误差只增长了 0.037, 故此选择岭参数为 0.4 时, 进行回归分析. 得到回归方程如下:

$$Y = 11.9905 - 0.0134X_2 - 0.00092X_3 + 0.0735X_5 + 0.00059X_6 - 0.10201X_{10} - 0.06712X_{11} \pm 4.218$$

式中,  $Y$  代表预测的 APN 参考值, 4.218 为剩余标准差.

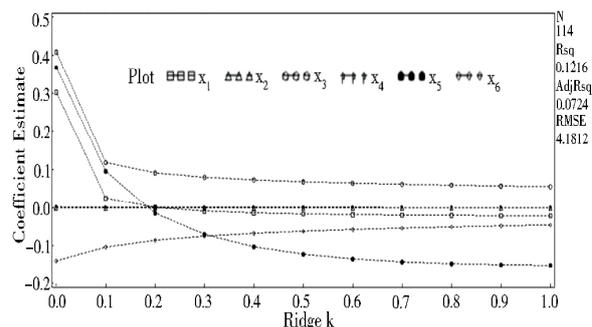


图 2 APN 岭迹图

Fig.2 Ridge trace of APN

### 2.6 支持向量机模型

在 Clementine12.0 软件中应用 SVM 模块建立数据流, 以 APN 参考值为输出变量, 相关的 6 项地理指标为输入变量, 分别选择不同的核函数

进行训练,并给出不同核函数对应的 SVM 模型的均方误差(MSE)和平均绝对误差(MAE),结果见表 3.根据表 3 可知,以 RBF 为核函数构建的 SVM 模型预测精度较好,因此选择 RBF 核函数的 SVM 模型作为 APN 参考值的预测模型.

表 3 各核函数的预测误差

Table 3 Prediction error of each kernel function

核函数	MSE	MAE
RBF	0.282	2.765
Polynomial	0.477	2.494
Sigmoid	1.442	6.779
Linear	0.327	2.745

### 2.7 模型质量评价

对岭回归和支持向量机预测模型进行模型质

表 4 APN 模型结果比较

Table 4 The models' result of APN

地点	实测值	岭回归预测	相对误差	<i>P</i>	支持向量机预测	相对误差	<i>P</i>
沈阳	11.08	10.93	1%	0.708	10.96	1%	0.592
南通	12.37	11.38	9%		10.87	14%	
柳州	13.37	13.06	2%		12.53	7%	
响水县	10.38	10.74	5%		10.38	8%	
唐山	11.44	10.77	6%		10.49	9%	
合肥	12.77	12.14	5%		11.92	7%	
百色	13.15	13.07	1%		12.50	5%	
上海	11.38	11.24	1%		10.76	6%	
湛江	11.75	11.75	0%		11.06	6%	
新乡	10.65	10.67	0%		10.32	3%	

量评价,选取部分城市的实测值与岭回归预测值和支持向量机预测值进行对比,结果见表 4.从表中可以看出,岭回归预测值的相对误差范围在 0%~9%之间,支持向量机预测值的相对误差范围在 1%~14%之间,岭回归预测值的误差相对较小.对两组预测值和实测值进行配对样本 *T* 检验后,两组预测值的 *P* 值均大于 0.05,可认为在 95%的显著水平下,预测值与实测值不存在显著差异,而岭回归预测值经 *T* 检验, $P=0.708$ ,大于支持向量机预测值的 *P* 值(0.592),说明岭回归方法预测精度更高,所以选择岭回归方法进行预测,预测全国 2 322 个数据点的 APN 参考值.

### 2.8 趋势分析结果

利用 ArcGIS 软件中地统计模块下的趋势面分析,对 APN 预测数据进行空间趋势观察.以东西方向为 X 轴,南北方向为 Y 轴,APN 预测数据为 Z 轴建立三维网格,将 2 322 个预测数据点投影到三维空间,进行 APN 预测值的趋势面分析,如图 3.从图中可以看出 APN 预测值在东西方向上(X 轴),由西向东呈明显递增趋势,拟合曲线呈非线性变化;在南北方向上(Y 轴),由北向南呈现先略微降低而后迅速上升的趋势,拟合曲线呈明显的下凹型,存在非线性变化.

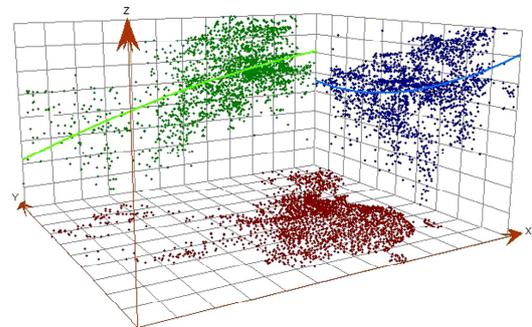


图 3 APN 趋势面分析

Fig.3 Trend surface analysis of APN

### 3 讨 论

根据相关分析结果, APN 参考值与 6 项地理指标存在相关性, 其中与  $X_3$  的相关性十分显著. 通过地理探测器的结果可知,  $X_3$  是最主要的影响因子, 结合建模预测后的趋势分析, 可以明显看出我国健康人 APN 参考值整体上呈现东南高于西北, 沿海高于内陆的分布趋势. 这与相关分析中, 与  $X_3$  呈负相关, 与  $X_5$  和  $X_6$  呈正相关的结果一致.

#### 3.1 海拔高度

我国地势西高东低, 自西向东呈阶梯状降低, 青藏高原更是有“世界屋脊”之称, 平均海拔在 4 000 m 以上. 随着海拔高度的增加, 空气变得稀薄, 空气中的氧气量也随之降低. 研究表明<sup>[20-21]</sup>, 低氧环境对脂肪细胞分泌因子(adipokines, ADS)会产生调节作用, 低氧环境通过激活低氧诱导因子 1(hypoxia-inducible factor-1, HIF-1)而降低 APN 的表达, 诱导脂肪细胞内活性氧自由基(reactive oxygen species, ROS)增加, 而 ROS 导致脂肪细胞中的 APN 的表达抑制, 减少了 APN 的合成, 导致高海拔低氧环境下, 人体内 APN 较低的现象.

#### 3.2 气候因素

气候因素的影响主要表现在气温和降水两方面, 我国气候类型多样, 水热条件组合复杂, 对当地环境产生不同的影响, 进而影响到人的生活方式及生活习惯等. 研究发现<sup>[22]</sup>, 在高温 33 °C、60%湿度下进行运动, 可升高 APN 水平. 我国南方地区属于亚热带季风气候, 水热同期, 夏季高温多雨, 湿度大, 长期在高温潮湿的环境下生活工作, APN 水平较其他地区有所上升, 这与本文研究结论一致. 不同的气候环境也会影响不同地区人群的膳食结构, 寒冷地区的人群由于环境的影响, 膳食结构中脂类物质摄入较多, 高脂饮食导致能量摄入过剩且体脂增加, APN 水平表达降低, 直接或间接地对健康造成影响<sup>[23]</sup>.

#### 3.3 土壤因素

土壤因素的影响主要是通过农作物间接影响人体健康, 不同地区土壤类型不同, 农作物种类繁

多, 南北方差异较大. 从相关分析的结果来看, APN 参考值与表层土壤 pH 和表层土壤总可交换量呈负相关关系, 我国土壤化学性质, 整体上存在“南酸北碱”的特性, 南方土壤偏酸性, 北方土壤偏碱性, 导致南北方农作物的差异. 东南沿海地区就西北内陆地区而言, 可摄入的高纤维素膳食和鱼油等较多, 膳食结构更为多元均衡, 而这些物质都能提高 APN 水平, 影响其表达<sup>[24]</sup>.

综上所述, 本研究可得到以下结论:

- (1) 地理环境对人体机能的影响存在于方方面面, 不同的地理环境造成了不同地区人群 APN 参考值的差异, 在空间上存在相关性.
- (2) 我国健康人 APN 参考值受海拔高度、气候条件和土壤因素的共同影响, 在空间上呈现东南高于西北, 沿海高于内陆的特征.
- (3) APN 参考值由基因和环境共同决定, 目前的医学参考值的制定忽略了地理环境的影响, 影响了医学研究临床判断的准确性. 研究地理环境对 APN 参考值的影响作用具有重要的科学意义. 若已知某一地区的纬度、海拔高度、年平均气温、年降水量、表土 pH 和表土总可交换量, 即可根据岭回归的预测公式对该地区健康人 APN 参考值进行预测. 例如西安地区的纬度 34.28°, 海拔 397.5 m, 年平均气温 13.7 °C, 年降水量 553.3 mm, 表层土壤 pH 值 7.8, 表层土壤总可交换量为 24.1 cmol/kg, 根据公式即可得出西安地区健康人 APN 参考值为  $(10.09 \pm 4.218)$  mg/L.

#### 参 考 文 献:

- [1] 刘维英. 基于 OSAHS 的间歇低氧对脂肪细胞分泌功能影响机制的研究[D]. 兰州: 兰州大学, 2014.
- [2] 黎丹凤, 张进安, 刘军. 2 型糖尿病患者糖尿病视网膜病变和脂联素的相关性分析[J]. 复旦学报: 医学版, 2016, 43(4): 453-456.
- [3] 刘岩, 邹大进, 李慧, 等. 低脂联素血症是冠状动脉粥样硬化严重程度的重要标志[J]. 中华内分泌代谢杂志, 2005, 21(1): 9-12.
- [4] 王鸿燕. 脂联素与胰岛素抵抗及血脂代谢关系的研究[D]. 上海: 第二军医大学, 2004.
- [5] 董靖, 李志勇, 卢旻鹏, 等. 西南地区人群尿脂联素水平与 2 型糖尿病合并冠心病的相关性[J]. 中国老年学杂志, 2017, 37(15): 3 763-3 765.
- [6] 白桦. 老年和青年不同 2 型糖尿病发病年龄与血清脂联素的关系[D]. 沈阳: 中国医科大学, 2007.

- [7] PUJEN T, LIN M L, CHU C M, et al. Spatial autocorrelation analysis of health care hotspots in Taiwan in 2006 [J]. BMC Public Health, 2009, 9 (1): 464.
- [8] 林锦耀, 黎夏. 基于空间自相关的东莞主体功能区划分[J]. 地理研究, 2014, 33(2): 349-357.
- [9] 王劲峰, 廖一兰, 刘鑫. 空间数据分析教程[M]. 北京: 科学出版社, 2010.
- [10] 王劲峰, 徐成东. 地理探测器: 原理与展望[J]. 地理学报, 2017, 72(1): 1 805-1 814.
- [11] WANG X G, XI J C, YANG D Y, et al. Spatial differentiation of rural touristization and its determinants in China; a geo-detector-based case study of Yesanpo scenic area [J]. J. Resour. Ecol., 2016, 7 (6): 464-471.
- [12] 李晓龙. 我国流行性乙型脑炎时空分布特征及相关地理气候因素研究[D]. 北京: 中国疾病预防控制中心, 2014.
- [13] 宇传华. SPSS 与统计分析[M]. 北京: 电子工业出版社, 2007.
- [14] 李鹏. 健康成年人两项肝功能指标空间分布特征[D]. 西安: 陕西师范大学, 2018.
- [15] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2-10.
- [16] 李成思. 基于相对误差意义下的最小二乘法[J]. 数理统计与管理, 2003, 22(4): 36-40.
- [17] 李孟姣, 葛森, 王聪霞, 等. 基于神经网络的血浆纤维蛋白原参考值与地理因素的关系[J]. 南方医科大学学报, 2016, 36(8): 1 062-1 068.
- [18] CHANG K T. 地理信息系统导论(原著第 8 版)[M]. 陈健飞, 译. 北京: 科学出版社, 2016.
- [19] 郭旭东, 傅伯杰, 马克明, 等. 基于 GIS 和地统计学的土壤养分空间变异特征研究——以河北省遵化市为例[J]. 应用生态学报, 2000, 11(4): 557-563.
- [20] 陈宝莹, 魏经国, 王玮, 等. 氧自由基对脂肪细胞中脂联素表达的下调作用[J]. 心脏杂志, 2008, 20(2): 142-146.
- [21] 陈宝莹. 低氧及氧化应激对脂肪细胞分泌因子表达的调节及其机制的研究[D]. 西安: 第四军医大学, 2006.
- [22] 许杰. 不同湿度高温环境有氧运动对男子摔跤运动员减体重的对比研究[D]. 北京: 北京体育大学, 2012.
- [23] 胡健翔. 膳食脂肪摄入与肥胖相关指标、血压的关系研究[D]. 北京: 中国疾病预防控制中心, 2017.
- [24] KASIM-KARAKAS S E, TSODIKOV A, SINGH U, et al. Responses of inflammatory markers to a low-fat, high-carbohydrate diet: effects of energy intake[J]. American Journal Clinical Nutrition, 2006 (4): 774-779.

## The Influence of Geographical Environment on Serum Adiponectin Levels in Healthy People

Ji Xu<sup>1</sup>, GE Miao<sup>1</sup>, LI Xiao-ping<sup>1</sup>, JING Jing<sup>1,2</sup>, LIN Ning<sup>1</sup>

(1. School of Geography and Tourism, Shaanxi Normal University, Xi'an 710119, China; 2. Department of Geography, College of Geography and Environment, Baoji University of Arts and Sciences, Baoji 721016, China)

**Abstract:** The objective of this study is to analysis geographical distribution characteristics of serum adiponectin(APN) levels for healthy people and explore the influence of geographical environment on it. A total of 139 88 samples of APN reference values from 106 administrative units were collected to study their correlation with geographical environment via the spatial autocorrelation analysis, the correlation analysis and the Geodetector. After comparing the ridge regression analysis and the support vector regression, the optimal model was chose. Spatial analysis of APN prediction data by using trend analysis. The result show that there were 6 geographic factors, latitude, altitude, annual mean temperature, annual precipitation, topsoil pH and total exchangeable amount of topsoil, correlated significantly with the APN reference values. The regression equation of optimal was  $Y = 11.990 5 - 0.013 4X_2 - 0.000 92X_3 + 0.073 5X_5 + 0.000 59X_6 - 0.102 01X_{10} - 0.067 12X_{11} \pm 4.218$ . The overall trend of the APN levels is high in southeast, but low in northwest and high in coastal, but low in inland.

**Keywords:** Serum adiponectin; Geographical environment; Geodetector; Ridge regression; Support vector regression