

# Spatial distribution characteristics of tumor marker CA724 reference values in China

Jing Jing<sup>1,2</sup>  | Miao Ge<sup>1</sup> | Ziqi Yang<sup>2</sup> | Peng Li<sup>3</sup>

<sup>1</sup>Institute of Healthy Geography, College of Geography and Tourism, Shaanxi Normal University, Xi'an, Shaanxi, China

<sup>2</sup>Baoji University of Arts and Sciences, Baoji, Shaanxi, China

<sup>3</sup>Department of Epidemiology and Biostatistics, School of Public Health, Xi'an Jiaotong University Health Science Center, Xi'an, Shaanxi, China

## Correspondence

Miao Ge, Institute of Healthy Geography, School of Geography and Tourism, Shaanxi Normal University, 620 West Chang'an Road, Xi'an, 710119 Shaanxi, China.  
Email: gemiao@snnu.edu.cn

## Funding information

Fundamental Research Funds For the Central Universities, Grant/Award Number: 2016TS055

## Abstract

**Objects:** This study aims to explore the Cancer antigen 724 (CA724) reference values spatial distribution characteristics in healthy Chinese adults. The study can provide regional reference for medical diagnosis.

**Study Design:** The relationship between CA724 and 25 geographical environmental factors was analyzed firstly. Artificial neural network simulation training was used to construct the prediction model. The national forecast distribution map of the CA724 reference values was obtained by the geostatistical mapping method. Analyzing and exploring the influence mechanism of geographical environment factors on CA724 reference values.

**Methods:** Collecting 34470 cases from more than 106 cities healthy adults CA724 reference values via several paper databases in 10 recent years. Correlation analysis, RBF artificial neural networks and trend surface analysis were applied to explore if there was any tendency of spatial variation. The Kriging interpolation of geostatistical analysis was developed to reveal the spatial distribution characteristics of the CA724 reference values.

**Results:** The distribution of CA724 reference values of Chinese healthy adults shows a downward trend from south to north. CA724 reference values have negative correlations with latitude, annual sunshine duration and topsoil cation exchange capacity in clay. CA724 have positive correlations with annual mean air temperature, annual mean relative humidity, and annual precipitation amount. High temperature and high humidity environment will reduce gastrointestinal function and breeze various mold bacteria. Lack of sunshine can easily lead to vitamin C deficiency in the body. These will increase the incidence of gastrointestinal diseases and gastric cancer, then increase the CA724 value.

**Conclusion:** CA724 reference values show spatial autocorrelation and regional variation. There are some geographical environment factors effected Chinese healthy adults CA724 reference values. Geographic factors such as sunshine, temperature, and humidity have effects on CA724 reference values can provide new ideas and directions of prevention and clinical diagnosis in the future.

**KEYWORDS**

CA724, gastric cancer, geographical environment factors, reference values, tumor markers

## 1 | INTRODUCTION

Cancer antigen 724 (CA724) is a high molecular weight mucin carcinoembryonic antigen and the molecular weight is >200 000. It has double antigenic determinant. Its level is related to tumor size, stage, and metastasis. CA724 is present in stomach, colon, pancreas, lung, and ovarian tumors for almost 85%-95%. However, it does not exist in benign tumors, exudates, or normal human tissues.<sup>1</sup> It is a broad-spectrum tumor marker. In patients with malignant tumors of the digestive system, serum is elevated. Gastric cancer patients have higher expression.<sup>2</sup> It is reported that CA724 has a sensitivity to gastrointestinal tumors of 78.8%, which is a reliable indicator for the diagnosis of gastric cancer.<sup>3-5</sup> Research reported that: serum CA724 level after gastric cancer surgery is significantly lower than before surgery, it can be used for monitoring the therapeutic effect.<sup>6</sup> The level of CA724 is closely related to clinical stage, tumor size, lymph node involvement and invasion and metastasis of gastric cancer, it has high specificity.<sup>7</sup> The positive rate of tumor markers in gastric cancer in all stages of gastric cancer is significantly correlated with tumor stage, prognosis, and depth of tumor invasion, and may indicate distant metastasis. It can also be used as an early monitoring indicator for recurrence after radical surgery and evaluate the chemotherapy effect of gastric cancer.<sup>8</sup> Obviously, the determination of the CA724 reference range is particularly important. Many factors affecting CA724 in existing research results include heredity, eating habits, lifestyle habits, sunshine, etc French researchers Martel et al<sup>9</sup> have presented some known risk factors, including genetic, dietary, behavioral, and *Helicobacter pylori* infection in gastric cancer. A research of Korea has pointed that salt or some chemicals contained in kimchi and soybean pastes, which are increased risk of gastric cancer.<sup>10</sup> You et al<sup>11</sup> concluded some influencing factors of gastric cancer such as *H. pylori* infection, cigarette smoking, and low levels of dietary vitamin C. In Japan, the negative correlation coefficients of male age adjusted stomach cancer mortality rate with yearly sunshine duration was significant, but no correlation was found with yearly average temperature. Calcium has antiphlogistic effect, because of its capillary permeability reducing action, which may also prevent lodging of tumor cells in hematogenous metastasis of cancer in microcirculation system. The sunshine duration may be concerned with calcium absorption through its action of vitamin D production on skin.<sup>12</sup> An investigation in Jiangsu province of China, risk factors for gastric cancer include family history of cancer, fast eating, eating salty food,

hot food, male smoking, etc<sup>13</sup> The CA724 shows higher value in patients with gastric cancer.

The human and geographical environment, however, is an organic and interactive whole. A lack of studies exists for the relationship between CA724 reference values and the geographical environment. This study focuses on exploring the relationship between CA724 reference values and geographical environmental factors, revealing the spatial characteristics of CA724 reference values in China. It can provide scientific basis for clinical diagnosis and treatment. This study provides a regional basis for clinical diagnosis and treatment.

## 2 | MATERIALS AND METHODS

### 2.1 | Study area and subjects

The 34 470 cases of Chinese healthy adults with CA724 reference values were collected. Organizing CA724 reference values from more than 190 administrative units in 106 cities and regions via several paper databases (2008-2018) (<http://www.cnki.net>). The cities are selected randomly. Selecting the units according to the unified sampling method. Western cities have less data than the eastern ones. There are 115 units from eastern regions, 42 units from middle regions, and 33 units from western regions due to disparities of population density. However, Hong Kong Special Administrative Region, Macao Special Administrative Region and Taiwan province were not involved in this study. The ratio of males (18075) to females (16375) was 1:1.1. The subjects were all apparently healthy with the inclusion criteria being normal weight and height, that is, (body mass index [BMI] between 18.5 and 25.0 kg/m<sup>2</sup>), normal blood pressure and normal blood glucose after 8 hours' fasting. In addition, *t* test was employed to test the disparities of CA724 reference values among different regions. *T* test result is  $P \frac{1}{4} 0.001 < 0.05$ , which indicate there is a significant difference of CA724 reference values among different regions. Serum samples were selected from healthy individuals in each unit during the same period to exclude gastric cancer or other diseases that could cause changes in CA724.

### 2.2 | Geographical Environment variables

Information on 4 classes of geographical environment factors were collected concluding spatial indicators, topographical indicator, meteorological indicators, and soil indicators. Spatial indicators contain Longitude ( $X_1$ ) and Latitude ( $X_2$ ). Topographical indicator contains Altitude

( $X_3$ ). Spatial location and terrain indices from relevant maps and dictionaries. Meteorological indicators contain 6 indexes ( $X_4$ - $X_9$ ). Meteorological index provided by the meteorological data sharing service system in China (<http://www.cma.gov.cn>). Soil indicators contain 14 indexes ( $X_{10}$ - $X_{25}$ ). Soil indexes obtained from the Harmonious World Soil Database (HWSD) of Vienna International Institute of Applied Systems (IIASA) supplied by the Food and Agriculture Organization (FAO) (<http://www.fao.org/nr/land/soils/harmonized-world-soil-database/zh/>). In all, 25 different geographic locations and environmental variables (referred to by abbreviations  $X_1$ - $X_{25}$ ) were used as shown in Table 1.

### 3 | METHODS

#### 3.1 | Correlation analysis

According to the principle and method of correlation analysis, the correlation between CA724 reference values and 25 geographical factors were analyzed. The principle analysis can find out the relevant geographical factors. The study was conducted using SPSS 22.0 software by comparing CA724 reference values and geographical environment factors. The correlation was considered statistically significant when the  $P$ -value was at the 95% confidence interval, and it was marked as \*. The correlation was considered as strongly statistically significant when the  $P$ -value achieved a 99% confidence interval  $t$ , and it was marked as \*\*.

#### 3.2 | Artificial neural networks

Artificial neural network is based on the structure of the brain tissue. The mechanism of understanding based on the

structure and behavior of an engineering system.<sup>14,15</sup> In the prediction function by the scientific community recognized as superior to the traditional statistical prediction method. It has a better application of expression in medical clinical diagnosis, medical imaging, prognosis, and clinical decision-making analysis and so on.

This study uses radial basis function neural network (RBFANN) for modeling. The basic idea of RBFANN is to use RBF as a function of hidden nodes. When the input layer directly passes signals to the hidden layer, the input vector can be directly mapped to the hidden space without connecting through the weights. Thus making the hidden layer node output.<sup>16</sup>  $n_i$  is represented the number of input layer nodes of the 3-layer forward neural network.  $n_h$  is expressed the number of hidden layer nodes.  $n_o$  is represented the number of output layer nodes. In order to find a more suitable number of hidden layer nodes in a large number of experiments and practical applications, the range of ideal hidden layer nodes appears to be:

$$\frac{n_i + n_o}{2} \leq n_h \leq (n_i + n_o) + 10 \quad (1)$$

After confirming the range of the number of hidden layer nodes, the RBFANN corresponding to the number of hidden layer nodes in the range is trained. After that training the same sample and comparing the test results.

Whether the distribution of CA724 reference values in this study are linear or nonlinear, relying on statistical methods or human experience to determine the results are not necessarily realistic. Therefore, the neural network is used to compare the linear model with the nonlinear model in order to predict the nonlinear relationship between the CA724 reference values and the geographical environment. RBFANN is used to build

**TABLE 1** Geographical environment factors

$X_n$	Geographical environment factors	$X_n$	Geographical environment factors	$X_n$	Geographical environment factors
$X_1$	Longitude (°)	$X_{10}$	Topsoil sand percentage (% wt.)	$X_{19}$	topsoil cation exchange capacity in silt (cmol/kg)
$X_2$	Latitude (°)	$X_{11}$	Topsoil powder percentage (% wt.)	$X_{20}$	Topsoil base saturation (%)
$X_3$	Altitude (m)	$X_{12}$	The percentage by overburden clay (% wt.)	$X_{21}$	Topsoil exchangeable base (cmol/kg)
$X_4$	Annual sunshine duration (h)	$X_{13}$	Topsoil reference capacity (kg/dm <sup>3</sup> )	$X_{22}$	Topsoil carbonate (lime) content (cmol/kg)
$X_5$	Annual mean temperature (°C)	$X_{14}$	Overburden soil bulk density (kg/dm <sup>3</sup> )	$X_{23}$	Topsoil sulfate (gypsum) content (cmol/kg)
$X_6$	Annual mean relative humidity (%)	$X_{15}$	Topsoil fertile volume percent (% vol.)	$X_{24}$	Surface soil alkalinity (cmol/kg)
$X_7$	Annual precipitation amount (mm)	$X_{16}$	Topsoil organic carbon (% wt.)	$X_{25}$	topsoil salinity (dS/m)
$X_8$	Annual range of temperature (°C)	$X_{17}$	Topsoil PH		
$X_9$	Annual mean wind speed (m/s)	$X_{18}$	Topsoil cation exchange capacity in clay (cmol/kg)		

network models. The reference values of CA724 are predicted by the geographical environment factors which have correlations with CA724.

### 3.3 | Spatial statistics analysis

Spatial statistical analysis reveals spatial changes and features through the distribution of data sets characterized by geographic attributes. The core concept is derived from Tobler's first law of geography. Geographic phenomena or objects exhibit spatial dependence and variation characteristics.<sup>17</sup> Under the long-term influence of the geographical environment, the adaptation of bodily functions has led to various regional characteristics.

### 3.4 | Spatial autocorrelation

Spatial autocorrelation is mainly used to analyze the correlation and dependence between geographical phenomena or attributes and their adjacent units in a certain place. Its purpose is to reveal the dependence.<sup>18,19</sup> Moran's *I* is the main indicators for assessing spatial dependence and correlation.<sup>20,21</sup> We used Moran's *I* index to assess the degree of dependence and correlation of CA724 reference values in space. The threshold range of Moran's *I* is  $[-1, 1]$ . When Moran's *I* index is greater than zero, it represents a positive statistical correlation between spatial objects. Conversely, when it is less than zero, there is a negative correlation. When the index is the same as zero, the mean shows the balanced distribution of the spatial objects.<sup>22,23</sup>

The significance of the Moran's *I* index is tested by the standard statistic-*Z* value. We suggest that the determination of significance depends on testing the *Z* value with the *P* value. It is statistically significant if the *Z* value is greater than zero. However, it indicates a negative statistical significance if the *Z* value is less than zero. When the *Z* value is equal to zero, it presents a random distribution.

### 3.5 | Trend surface analysis

Trend surface analysis is a multivariate analysis of data fitted by abstract mathematical curves based on least squares.<sup>24</sup> In general, this surface includes trend surface and residual surface. The trend surface mainly reflects the macroscopic distribution of geographic datasets, while the residual surface mainly reflects the influence of random factors on the overall distribution of microscopic space. Therefore, it is necessary to reduce the residual value as soon as possible to eliminate the influence of random variables. The purpose of the random variable reflects the spatial distribution of the sample and improves the simulation accuracy.<sup>25</sup> Using trend surface analysis to analyze the global trend of CA724 reference prediction data and establish a three-dimensional view. Each point in the data set is projected on

an orthogonal plane in the east-west and north-south directions. Then based on the corresponding projections, the best fit lines are determined and the corresponding presence trends are described by it.

### 3.6 | Spatial distribution

In order to reveal the distribution law of Chinese adults' CA724 reference values, 2322 cities in China were selected as the observation points. The predicted values of CA724 were calculated. The Kriging interpolation was performed by using the spatial analysis module of ArcGIS software and the spatial distribution map was output.<sup>26,27</sup> Kriging interpolation method considers the shape information of the sample point, the spatial orientation, and the structural information provided by the variogram. Then estimating the prediction data unbiased based on the sample data of CA724. Kriging interpolation can be used to obtain the spatial distribution trend of the CA724 reference values in China.

## 4 | RESULTS

### 4.1 | Correlation analysis

The correlations between CA724 and geographical environment factors are shown in Table 2. It shows that the CA724 is significantly correlated with geographical environment factors in China. Apparently, the correlation between CA724 and  $X_2, X_5, X_7, X_{18}$  was significant ( $P = 0.015, 0.037, 0.048, 0.040, <0.05$ ); the correlation between CA724 and  $X_4, X_6$  was very significant ( $P = 0.003, 0.008, <0.01$ ). The CA724 has a negative correlation with Latitude ( $X_2$ ), Annual sunshine duration ( $X_4$ ), and topsoil cation exchange capacity in clay ( $X_{18}$ ), but has a positive correlation with annual mean air temperature ( $X_5$ ), annual mean relative humidity ( $X_6$ ) and annual precipitation amount ( $X_7$ ).

### 4.2 | Artificial neural networks

According to formula (1), the number of geographic factors associated with the CA724 reference values is 6, the corresponding input layer node number is  $n_i = 6$ , and the number of output layer nodes is  $n_o = 1$ , so the corresponding hidden layer node number  $n_h$  is in the range of  $[4, 17]$ .

Importing the training data into SPSS Clementine 12.0 and establish a data stream. Connecting the neural net node in the modeling module to the appropriate location of the data stream. Establishing the RBFANN model for each node number to repeatedly train the same sample data. The test samples are predicted by each model. The analysis node is added to the prediction result node to analyze the prediction results. (Table 3) Calculating the Mean Absolute Error (MAE) and Mean Squared Error (MSE) of the RBFANN model corresponding to the number of hidden

**TABLE 2** Correlation and significance of CA724 in healthy Chinese adults

Geographical factors ( $X_n$ )	$t$	$P$
$X_1$	-0.025	0.733
$X_2$	-0.176	0.015*
$X_3$	0.028	0.699
$X_4$	-0.217	0.003**
$X_5$	0.151	0.037*
$X_6$	0.193	0.008**
$X_7$	0.144	0.048*
$X_8$	-0.124	0.088
$X_9$	-0.049	0.499
$X_{10}$	-0.076	0.298
$X_{11}$	0.106	0.145
$X_{12}$	-0.024	0.738
$X_{13}$	-0.024	0.744
$X_{14}$	-0.016	0.822
$X_{15}$	0.117	0.108
$X_{16}$	0.026	0.724
$X_{17}$	0.074	0.310
$X_{18}$	-0.149	0.040*
$X_{19}$	-0.100	0.172
$X_{20}$	0.062	0.398
$X_{21}$	0.078	0.286
$X_{22}$	0.066	0.363
$X_{23}$	-0.100	0.169
$X_{24}$	-0.049	0.500
$X_{25}$	0.095	0.151

\*Correlation is significant at the 0.05 level (2-tailed).

\*\*Correlation is significant at the 0.01 level (2-tailed).

nodes. According to the results, when the number of hidden layer nodes is 5, the prediction effect is best. Therefore, the model corresponding to the number of five hidden layer nodes is selected as the RBFANN prediction model of the CA724 reference values.

### 4.3 | Spatial autocorrelation

As shown in Figure 1 and Table 4, Moran's  $I$  was equal to 2.301783; the  $Z$ -value was 20.934111 by the test of auto-correlation. The result was greater than 2.58 in the 99% confidence interval. The  $P$ -value <0.01. In addition, Figure 1 showed a significant autocorrelation of the CA724 reference values in spatial distribution. It is indicated that the spatial disparities generated the regional distribution of the reference values of CA724. CA724 reference values were influenced by spatial disparities and geographical factors.

**TABLE 3** Mean Absolute Error (MAE) and Mean Squared Error (MSE) of radial basis function neural networks with different node of hidden layer

Node of hidden layer	MAE	MSE
4	1.075	1.411
5	1.046	1.379
6	1.052	1.396
7	1.087	1.405
8	1.071	1.392
9	1.068	1.388
10	1.073	1.388
11	1.044	1.389
12	1.101	1.424
13	1.076	1.438
14	1.110	1.430
15	1.123	1.468
16	1.066	1.412
17	1.088	1.454

### 4.4 | Trend surface analysis

A trend surface analysis was developed to explore the spatial trend of the CA724 reference values. We can clearly see 2 curves reflecting the reference values of CA724 in a 3-dimensional grid. The red curve was designated as the direction of from west to east, and the blue curve was designated as the direction from south to north. The CA724 reference values are gradually reduced from south to north, which was consistent with the spatial distribution. Speculatively, there is a distinctly regional distribution in space.

As presented in Figure 2, the CA724 values for healthy adults along the longitude it first increased from west to east and then decreased slightly in the extreme (the red curve), while it showed a decreasing trend from south to north along the latitude (the blue curve). With the ordinate ( $x$  axis) depicting longitude, the abscissa ( $y$  direction) latitude and  $z$  the CA724 concentration, it can be seen that the CA724 reference values first increase in the western direction along the longitude and then decrease in the eastern direction at the end (the red slope), while they decrease in the northern direction along the latitude (the blue curve).

### 4.5 | Spatial distribution

In order to clearly and accurately show the geographical distribution law of the CA724 of Chinese healthy adults. Using geostatistics module in ArcGIS software, 2322 Chinese cities were located in Chinese map. The Kriging interpolation was conducted after analyzing, screening, and transforming the data which were predicted by the optimal predictive model.

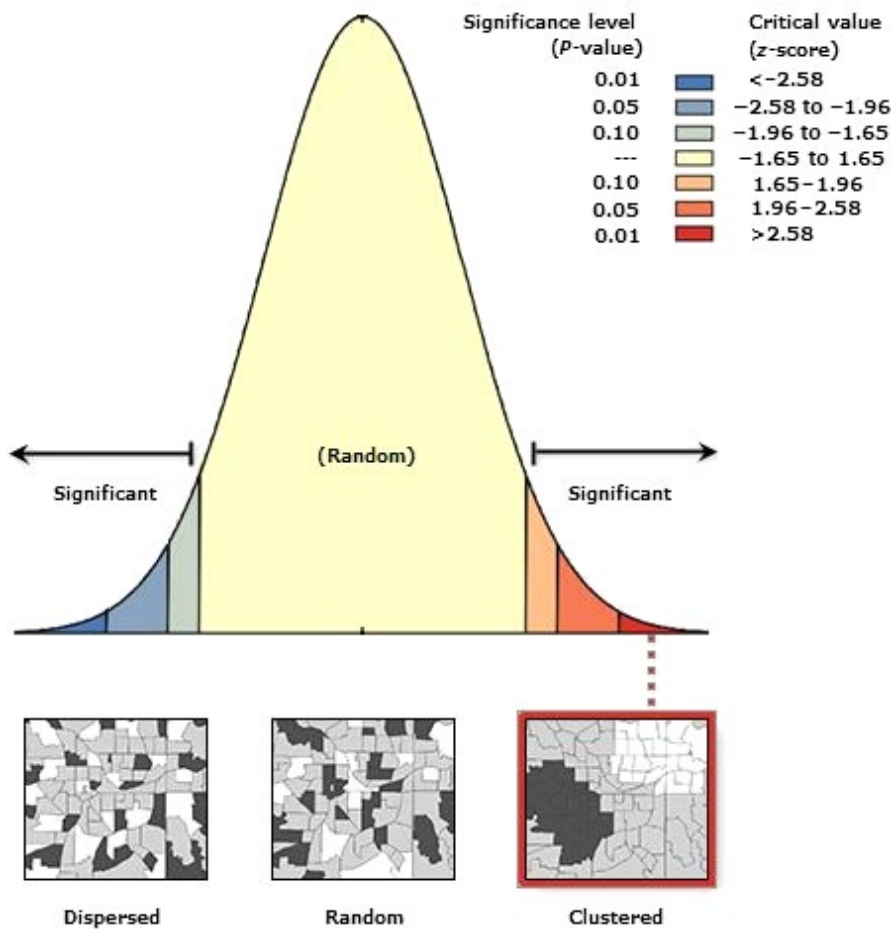


FIGURE 1 The spatial distribution of CA724 is clustered or dispersed

TABLE 4 Spatial autocorrelation report—Moran's results

Moran I	Expected	Index variance	Z-value	P-value
2.301783	-0.007299	0.012167	20.934111	0.000000

The spatial tendency chart of the Chinese healthy adults' CA724 reference values was output (Figure 3).

It can be seen from Figure 3, the distribution of Chinese healthy adults' CA724 reference values shows a downward trend from south to north. CA724 reference values have negative correlations with latitude ( $X_2$ ) annual sunshine duration

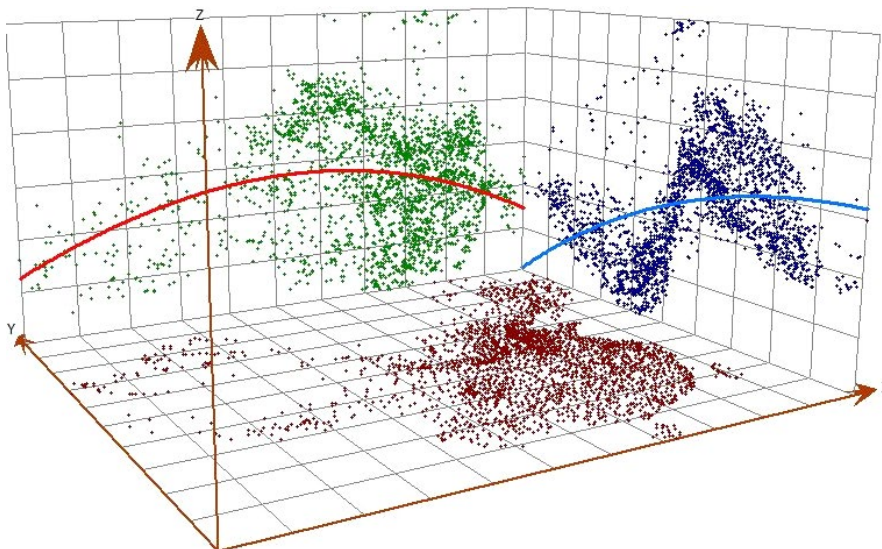
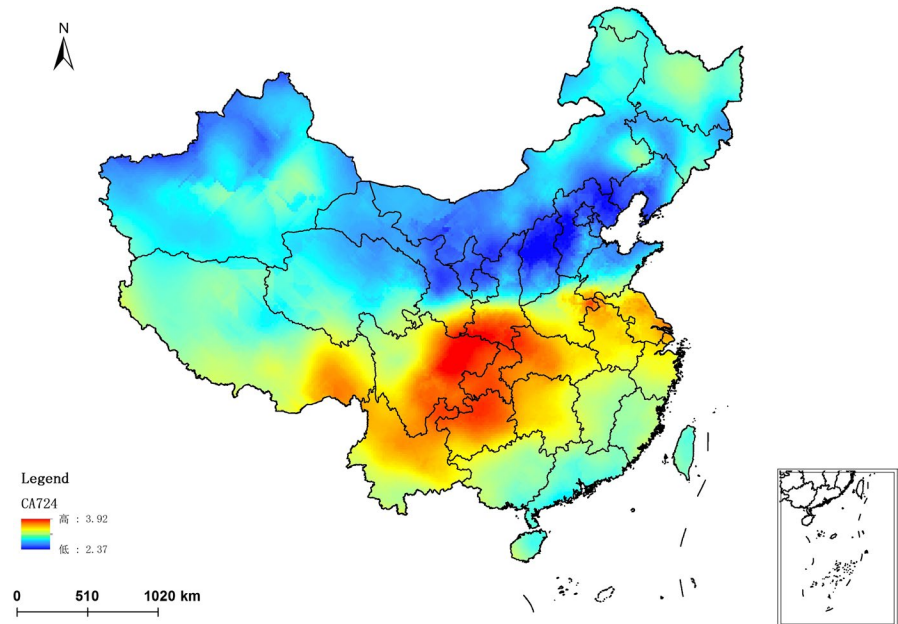


FIGURE 2 CA724 in healthy Chinese adults trend analysis chart. Red line is from West to East, Blue line is from North to South.

**FIGURE 3** Spatial tendency chart of the reference values for CA724 of Chinese healthy adults



( $X_4$ ) and topsoil cation exchange capacity in clay ( $X_{18}$ ). They have positive correlations with annual mean air temperature ( $X_5$ ), annual mean relative humidity ( $X_6$ ), and annual precipitation amount ( $X_7$ ). The trend consistent with the distribution law of annual mean relative humidity and in contrary to annual sunshine duration distribution of China. With the increase of latitude and annual sunshine duration, the CA724 reference values of Chinese healthy adults decreased. With the increase of annual mean air temperature ( $X_5$ ), annual mean relative humidity ( $X_6$ ) and annual precipitation amount, the CA724 reference values increased in Chinese healthy adults. This shows that Chinese healthy adults CA724 reference values and geographical environment factors have a correlation in spatial distribution.

## 5 | DISCUSSION

From the results of this study we can see that geographical environment factors relatively relevant with the Chinese healthy adults' CA724 reference values including latitude, annual sunshine duration, annual mean air temperature, annual mean relative humidity, annual precipitation amount, and topsoil cation exchange capacity in clay. The CA724 reference values have negative correlations with latitude, annual sunshine duration, and topsoil cation exchange capacity in clay, but have positive correlations with annual mean air temperature, annual mean relative humidity and annual precipitation amount.

### 5.1 | The influence of temperature for CA724

In epidemiology, the incidence of digestive ulcers increases from north to south. This is consistent with the findings

of this paper. Researchers of Japan have acquired data for infectious gastroenteritis cases and weather variability in Fukuoka, from 1999 to 2007 and used time-series analysis to assess the effects of weather variability on infectious gastroenteritis cases, adjusting for confounding factors. In total, 422 176 Infectious gastroenteritis cases were reported during the 9-year study period. The weekly number of infectious gastroenteritis cases increased by 7.7% (95% CI 4.6-10.8) for every 1°C increase in the average temperature.<sup>28</sup> A few studies in high-income countries have investigated the relationship between ambient temperature or precipitation and the occurrence of gastroenteritis. In most of the cases, hot temperatures and heavy precipitation events have been related to increases in infections. The findings suggest an important role of ambient temperatures, especially hot temperatures, in increasing gastroenteritis hospitalizations.<sup>29</sup> Various studies have shown that the incidence of gastrointestinal diseases increases with increasing temperature. High temperature conditions will reduce the secretion of digestive juices such as saliva, gastric juice, pancreatic juice, bile, and intestinal fluid. The digestive juices contain substances necessary for the digestion process of the food, such as free hydrochloric acid, protease, lipase, amylase, etc These substances can reduce the digestive function of the stomach. In high temperature environment, the increase of sweating causes a large loss of sodium chloride, and because the chloride ion of free hydrochloric acid in gastric juice comes from the blood, the acidity of gastric juice is also significantly reduced.<sup>30</sup> High temperature environment can also cause the gastric emptying to accelerate, so that the chemical digestion process of food can not be fully carried out. Because of the effects of temperature on the central nervous system of people,

high temperatures often lead to loss of appetite and irregular diet.<sup>31</sup> People eating melon and fruit cold drinks in high temperature weather increases the risk of gastrointestinal diseases. Affected by many factors above, high temperature can lead to high incidence of gastrointestinal diseases, resulting in an increase in CA724.

## 5.2 | The influence of humidity for CA724

The relative humidity of about 50% is the most unfavorable humidity condition for the disease of the digestive system. The threshold range corresponds to the most comfortable humidity range that the human body feels.<sup>32</sup> The annual average relative humidity in southern China just fits this range. Increased humidity causes food to become mildewed, and the growth of various molds greatly increases the incidence of gastric cancer. High temperature and high humidity are easy to breed various bacteria. Fungi such as *Aspergillus flavus*, *Bacillus cereus*, *Botox*, *Salmonella typhimurium*, etc are beneficial for growth and reproduction of digestive tract diseases. The fact that *Aspergillus flavus* is found in soya bean cakes raises the possibility that aflatoxins produced by this mold which grows on the soya bean cakes (from which soya bean paste is made) may be a possible etiologic factor in the high incidence of stomach cancer.<sup>33</sup> Infection of *Salmonella typhimurium* (*Salmonella typhi*) can lead to various organ diseases. This research first proposed that *Salmonella typhi*-infection could result in gastric oxidative stress and hemorrhagic ulcers.<sup>34</sup> Bacterial growth and reproduction are inseparable from high temperature and high humidity weather conditions, which can enable bacteria to grow and multiply better. In short, as the temperature and humidity increase, the proliferation of various molds increases, leading to an increase in the incidence of various gastrointestinal diseases, resulting in an increase in CA724. Therefore, the CA724 reference values of Chinese healthy adults are higher in the south than in the north.

## 5.3 | The influence of sunshine for CA724

Studies have shown that human body under the action of sun through the skin synthesis of large amounts of vitamin D. More than 80% of vitamin D from the sun in the ultraviolet radiation less through the source of nutrients provided, because vitamin D from the natural food are very few except fish. The factors that affect the skin synthesis of vitamin D are mainly due to adequate exposure to sunlight and dietary supplementation to achieve adequate levels of vitamin D.<sup>35,36</sup> Vitamin D not only maintains calcium and bone homeostasis, but also mostly inhibits tumor genesis, invasion, and metastasis through activation of vitamin D receptor. accumulating evidence from gastric cancer cells, animal models, and clinical trials suggest that vitamin D deficiency may increase the risk and mortality of gastric cancer.<sup>37</sup> Vitamin D-3

upregulated protein 1 (VDUP1) is a potent tumor suppressor whose expression is dramatically reduced in various types of human cancers, including gastric cancer.<sup>38</sup> Ample sunshine can provide the body with adequate vitamin D, which can inhibit the risk of stomach cancer. In contrast, fewer sunshine hours in the year will increase the incidence of *Helicobacter pylori* and gastric cancer, resulting in a high CA724 value.

## 5.4 | The influence of soil for CA724

Generally, the effect of soil on the human body is relatively indirect. It is usually caused by the influence of the soil on the crop, then the re-effect of the crop or food on the person. The soil cation exchange amount directly reflects the soil fertilizer retention, fertilizer supply performance and buffer capacity. In the case that the normal soil is not subjected to severe heavy metals or other pollution. The stronger soil topsoil clay cation exchange energy is the better soil fertility is. This will help plants absorb, purify water, and grow their crops healthily. A healthy crop or food will be beneficial to the health of the human gastrointestinal tract. So healthy conditions of people will show a low CA724 value.

## 6 | CONCLUSIONS

The present research provides a new view and insight in medical geography science through studying and exploring. The geographical environment and the human body are an organic integration. Various environments cause differences in our bodies, but CA724 reference values show similar and spatial dependence. We have probed the relationship between CA724 reference values of Chinese healthy adults and geographical environmental factors. In this paper, we choose the spatial regions of China as the research object, using the correlation analysis, spatial autocorrelation analysis, RBF Artificial neural networks, trend surface analysis, spatial analysis, and other methods to predict models. This study also explored the spatial distribution characteristics of CA724 reference values for clinical practice and future study. The laws of the distribution of Chinese healthy adults CA724 reference values shows the overall trend that a downward trend from south to north. There are some geographical environment factors effected Chinese healthy adults CA724 reference values. The spatial distribution of Chinese healthy adults CA724 reference values is consistent with latitude, annual sunshine duration, annual mean air temperature, annual mean relative humidity, annual precipitation amount, and topsoil cation exchange capacity in clay distribution.

Geographical environmental factors encompass a wide range of content. Many factors such as the ecological environment in which humans live, the water quality conditions of domestic water, the air quality in the living environment,



the soil conditions, and many geographical conditions affect the changes in the reference values of medical indicators. The geographical environmental factors selected in this paper are still not comprehensive. In the future research, we will also consider increasing geographical environment factors such as water environment, atmospheric environment, and human environment which have impacts on human health. It is currently known that changes in the geographical environment do have an impact on human health and physiological indicators. How the geographical environment factors affect and act on the human body remains to be further scientifically and deeply explained. In the future, we will further try to analyze or simulate the mechanism of influence mechanism by strengthen the professional study of medicine and geography knowledge. On the one hand the study can provide scientific reference for medical diagnosis, on the other hand it can enrich the contents of medical geography. In future studies, it is expected that the study will be further refined, in order to provide new ideas and scientific basis for the prevention, early detection and treatment of cancer such as gastric cancer.

## ACKNOWLEDGMENTS

The authors would like to thank all of the volunteers that took part in this study and the people for their assistance in technical and laboratory support. This study was by the Fundamental Research Funds For the Central Universities 2016TS055.

## CONFLICT OF INTEREST

None of the authors have any financial or other potential conflict of interest for this study.

## ETHICAL APPROVAL

The study was approved by Medical Ethics Committee of Second Affiliated Hospital of Xi'an Jiaotong University medical school (2010-LS-009).

## ORCID

Jing Jing  <https://orcid.org/0000-0002-5841-4826>

## REFERENCES

- Cheng J, Wang YQ. Combining testing gastric juice and serum levels of CA72-4, CA19-9 and CEA for the clinical value of gastric cancer. *J Chongqing Med Univ.* 2003;28(1):19-20.
- Moore AG, Brown DA, Fairlie WD, et al. The transforming growth factors super-family cytokine macrophage inhibitory cytokine-1 is present in high concentrations in the serum of pregnant women. *J Clin Endocrinol Metab.* 2000;85:4781-4788.
- Chen H. Clinical significance of combined detection of CA 724, CA 199 and CEA in gastro intestinal tract tumor. *J Taishan Med Col.* 2008;7(29):495-596.
- Mattar R, Andrade CRA, DiFavero GM, Gama-Rodrigues JJ, Laudanna AA. Preoperative serum levels of CA724, CEA, CA199 and alpha-fetoprotein in patients with gastric cancer. *Rev Hosp Clin Fac Med Sao Paulo.* 2002;57(3):89-92.
- Gaspar MJ, Arribas I, Coca MC, Díez-Alonso M. Prognostic value of carcino embryonic antigen CA199 and CA724 in gastric carcinoma. *Tumor Biol.* 2001;22(5):318-322.
- Yang H, Filipovic Z, Brown D, et al. Macrophage inhibitory cytokine-1: a novel biomarker for p53 pathway activation. *Mol Cancer Ther.* 2003;2:1023-1029.
- Takahashi Y, Takeuchi T, Sakamoto J, et al. The usefulness of CEA and/or CA199 in monitoring for recurrence in gastric cancer patients: a prospective clinical study. *Gas cancer.* 2003;6(3):142-145.
- Wang L. The significance of the combined measurement of serum MIC-1/CEA, CA724, CA50 in gastric cancer. *Qinghai Univ;* 2013.
- Martel CD, Forman D, Plummer M. Gastric cancer: epidemiology and risk factors. *Gastroenterol Clin North Am.* 2013;42(2):219.
- Nan HM, Park JW, Song YJ, et al. Kimchi and soybean pastes are risk factors of gastric cancer. *World J Gastroenterol.* 2005;11(21):3175-3181.
- You W, Zhang L, Yang C, et al. Gastric dysplasia and gastric cancer: *Helicobacter pylori*, serum vitamin C, and other risk factors. *J Natl Cancer Inst.* 2000;92(19):1607-1612.
- Takahashi E. Stomach cancer and ecologic factors in Japan. *Tohoku J Exp Med.* 1974;113(2):129-133.
- Liu AM. Epidemiological study on influencing factors of gastric cancer in high and low incidence areas of malignant tumors in Jiangsu Province. *Nanjing Med Univ;* 2007.
- Ge ZX, Sun ZQ. *Neural Network Theory and the Implementation of MATLAB R2007.* Beijing: Pub House of Elec Ind; 2007:110.
- Pattie DC, Snyder J. Using a neural network to forecast visitor behavior. *Ann Tourism Res.* 1996;23(1):151-164.
- Liu B, Guo H. *MATLAB Neural Network Super Learning Manual.* Beijing: People's Posts and Tele Press; 2014: 158: 310-316.
- Xu JH. *Quantitative Geography.* Beijing: Adv Edu Press; 2013.
- Bi SB. *Spatial Data Analysis.* Beijing: Peking Univ Press; 2015.
- Lin JX, Li X. The main functional division areas based on spatial autocorrelation: Dongguan. *Geo Res.* 2014;33(2):349-357.
- Gong YL, Zhang SL, Zhang LL. Spatial auto-correlation of urban land price: a case study of Suzhou. *Eco Geo.* 2011;31(11):1906-1911.
- Mao XC, Huang JX. *Spatial Analysis Modeling and Application.* Beijing: Sci Press; 2015.
- Wang HG, Wang WC, Peng B, et al. Tuberculosis incidence spatial analysis. *Univ Third Mil Med.* 2006;28(6):859-860.
- Zhang HH, Zeng YN, Jin XB. Spatial autocorrelation analysis of urban land price in Nanjing. *Nanjing Univ Technol Nat Sci.* 2009;45(6):821-830.
- Xu JH. *Mathematical Methods of Modern Geography.* Beijing: Advd Edu Press; 2009.
- Luo S, Ma L, Chen JW. Cancer mortality trends in the geographical distribution of surface analysis. *Chin Health Stat.* 2008;25(4): 357-360.

26. Tang G, Yang X. *ArcGIS Tutorial of GIS spatial Analysis*. Beijing: Sci Press; 2006:50-203.
27. Dang A, Jia HF, Yi SZ. *ArcGIS 8 Desktop Geographic Information System Application Guide*. Beijing: Tsinghua Uni Press; 2003: 79-301.
28. Onozuka D, Hashizume M, Hagihara A. Effects of weather variability on infectious gastroenteritis. *Epidemiol Infect*. 2010;2(138): 236-243.
29. Morral-Puigmal C, Martínez-Solanas È, Villanueva CM, Basagaña X. Weather and gastrointestinal disease in Spain: a retrospective time series regression study. *Env Int*. 2018;121(1):649-657.
30. Zhang JK. Digestive systemic disease. *Chi J Pra Int Med*. 1991; 1:5-5.
31. Parsons K. Human thermal environments: the effects of hot, moderate, and cold environments on human health, comfort and performance, Second Edition. *Safety Sci*. 1994;18(1):67-68.
32. Ma P, Ning GC, Zhao XY, et al. The lag effect of meteorological conditions on digestive diseases in Beijing. *J Lanzhou Uni (Nat Sci Edi)*. 2016;52(06):804-812,820.
33. Montesano R, Hall J. Environmental causes of human cancers. *Eur J Cancer*. 2001;37(suppl):67-87.
34. Hung CR, Wang PS. Gastric oxidative stress and hemorrhagic ulcer in *Salmonella typhimurium*-infected rats. *Eur J Phar*. 2004;491(1):61-68.
35. Akhtar S. Vitamin D status in South Asian populations—risks and opportunities. *Crit Rev Food Sci Nutr*. 2016;56(11):1925-1940.
36. Saraff V, Shaw N. Sunshine and vitamin D. *Arch Dis Child*. 2016;101(2):190-192.
37. Du C, Yang S, Zhao X, Dong H. Pathogenic roles of alterations in vitamin D and vitamin D receptor in gastric tumorigenesis. *Oncotarget*. 2017;8(17):29474-29486.
38. Kwon HJ, Won YS, Nam KT, et al. Vitamin D-3 upregulated protein 1 deficiency promotes N-methyl-N-nitrosourea and *Helicobacter pylori*-induced gastric carcinogenesis in mice. *Gut*. 2012;61(1):53-63.

**How to cite this article:** Jing J, Ge M, Yang Z, Li P. Spatial distribution characteristics of tumor marker CA724 reference values in China. *Cancer Med*. 2019;8:4465–4474. <https://doi.org/10.1002/cam4.2176>