Geographical Distribution of Serum Creatinine Reference Values in Healthy Adults: Postprint

Authors: Wei Dezhi, Ge Miao, Wang Congxia, Lin Qianyi, Li Mengjiao, Li Peng

Date: 2017-12-21T00:00:00+00:00

Abstract

Objective: To explore the relationship between serum creatinine (Scr) reference values and geographic factors in healthy adults, and to provide a scientific basis for establishing region-specific Scr reference value standards.

Methods: Scr reference values of 29,697 healthy adults measured by 347 medical institutions across 23 provinces, 4 municipalities, and 5 autonomous regions were collected. Twenty-three geographic data items were selected for correlation analysis with Scr reference values to identify geographic factors significantly associated with Scr reference values. Models were developed separately through principal component analysis and ridge regression analysis, the goodness of fit between measured and predicted values was compared to select the optimal prediction model, and finally a spatial distribution map of Scr reference values for healthy adults in China was constructed using Kriging interpolation.

Results: Scr reference values in healthy adults showed significant relationships with seven geographic indicators including latitude, annual sunshine hours, annual mean temperature, annual mean relative humidity, annual precipitation, annual temperature range, and topsoil (silt) cation exchange capacity. The distribution trend of reference values was higher in the south and lower in the north, exhibiting a relatively regular latitudinal variation.

Conclusion: Prediction of Scr reference values for healthy adults can be performed when geographic factor data for a specific region are known. Incorporating geographic factors into medical analysis facilitates the determination of region-specific medical reference values according to local conditions and improves the accuracy of clinical diagnosis.

Full Text

Preamble

Geographical Distribution of Serum Creatinine Reference Values in Healthy Adults

Dezhi Wei[1], Miao Ge[1], Congxia Wang[2], Qianyi Lin[1], Mengjiao Li[1], Peng Li[1]
[1]Institute of Health Geography, College of Tourism and Environment, Shaanxi Normal University, Xi' an 710119, China
[2]Xi' an Jiaotong University College of Medicine, Xi' an 710061, China

---

Abstract

Objective: To explore the relationship between serum creatinine (Scr) reference values in healthy adults and geographic factors, providing a scientific basis for establishing region-specific Scr reference value standards.

Methods: We collected 29,697 Scr reference values from healthy adults measured by 347 medical institutions across 23 provinces, 4 municipalities, and 5 autonomous regions in China. Twenty-three geographic variables were selected for correlation analysis with Scr reference values to identify significantly correlated factors. Predictive models were developed using principal component analysis and ridge regression analysis. The optimal model was selected by comparing the goodness-of-fit between predicted and measured values. Finally, the spatial distribution map of Scr reference values for healthy adults in China was generated using the Kriging interpolation method.

Results: Seven geographic factors demonstrated significant relationships with Scr reference values in healthy adults: latitude, annual sunshine duration, annual average temperature, annual average relative humidity, annual precipitation, annual temperature range, and topsoil (silt) cation exchange capacity. The reference values exhibited a clear distribution pattern, being higher in southern regions and lower in northern regions, showing consistent variation with latitude.

Conclusion: Knowing the geographic factors of a given region enables prediction of Scr reference values for healthy adults. Incorporating geographic factors into medical analysis facilitates the determination of appropriate medical reference values for different regions and improves the accuracy of clinical diagnosis.

Keywords: serum creatinine; geographical factors; principal component analysis; ridge regression analysis

---

Introduction

Creatinine is a metabolic product of creatine and phosphocreatine in muscle tissue, and its formation in the human body is relatively constant under normal conditions. Serum creatinine (Scr) is not only a mandatory item in health examinations but also one of the most widely used serum markers for kidney function assessment in clinical practice. It plays an important role in diagnosing kidney diseases, monitoring renal transplant rejection, and serving as an early indicator of kidney damage from chemotherapeutic drugs [1-3].

Currently, Scr reference values for healthy adults are applied rather generically, with identical reference values being used blindly across different geographic regions. The lack of region-specific Scr reference value standards for healthy adults has seriously compromised the accuracy of clinical diagnosis. Moreover, the relationship between Scr reference values and geographic factors in healthy adults has not been elucidated. To investigate this relationship and provide a scientific basis for establishing Scr reference values, this study employs ridge regression analysis and principal component analysis to examine the relationship between Scr reference values and regional geographic factors across China, developing two different predictive models and selecting the optimal one through comparison.

Methods

1.1 Geographic Environmental Indicators

This study selected four categories of natural geographic environmental indicators closely related to human health: geographic location, terrain, climate, and soil. Data for geographic location and terrain were obtained from the National Administration of Surveying, Mapping and Geoinformation Data Center, while climate data were sourced from the China Meteorological Data Sharing Service System and soil data from the World Soil Database. The geographic indicators included: X1 latitude (°), X2 longitude (°), and X3 altitude (m). Climate indicators comprised: X4 annual sunshine duration (h), X5 annual average temperature (°C), X6 annual average relative humidity (%), X7 annual precipitation (mm), X8 annual temperature range (°C), and X9 annual average wind speed (m/s). Soil indicators included: X10 percentage of topsoil gravel (%), X11 percentage of topsoil silt (%), X12 percentage of topsoil clay (%), X13 topsoil reference capacity (kg/dm^3), X14 topsoil bulk density (kg/dm^3), X15 topsoil gravel content (% vol.), X16 topsoil organic content (% wt.), X17 topsoil pH, X18 topsoil (clay) cation exchange capacity (cmol/kg), X19 topsoil (silt) cation exchange capacity (cmol/kg), X20 topsoil basic saturation (%), X21 topsoil total exchange capacity (cmol/kg), X22 topsoil alkalinity (cmol/kg), and X23 topsoil salinity (dS/m) [4-5]. These geographic indicators were matched with sample points based on regional names from the database to prepare for further data analysis.

1.2 Scr Reference Value Data

We retrieved 29,697 Scr reference values (mol/L) for healthy adults measured by 347 provincial, municipal, and county-level medical institutions and research organizations across 23 provinces, 4 municipalities, and 5 autonomous regions in China through searches of CNKI, CSCD, Wanfang, and VIP databases. The subjects ranged in age from 18 to 88 years. Data were more abundant for eastern plains than western plateau regions, with no data available for Hong Kong, Macau, or Taiwan [Figure 1: see original paper].

All selected Scr data were from healthy populations, with exclusion criteria including: dysfunction of major organs such as heart, liver, or kidneys; ketoacidosis; acute diabetes complications; trauma or surgery within 6 months; severe acute or chronic infections; use of nephrotoxic drugs; and primary hypertension. Measurements were performed using the sarcosine oxidase method and picric acid method, primarily with Olympus automatic biochemical analyzers, following referenced protocols [7-13].

1.3 Spatial Autocorrelation Analysis

Global spatial autocorrelation analysis was performed on the obtained data to determine the distribution pattern of sample points within the study area.

1.4 Correlation Analysis

To explore potential dependencies between Scr reference values in healthy adults and the selected geographic environmental indicators, multivariate correlation analysis was conducted between the medical and geographic variables.

1.5 Principal Component Analysis

Principal component analysis was applied to analyze various geographic factors, following methods described in referenced literature [13-15].

1.6 Ridge Regression Analysis

This method [16] was used to fit the relationship between Scr reference values and geographic factors in Chinese healthy adults, avoiding model failure due to multicollinearity.

1.7 Paired Sample T-Test

Following referenced methodology [17], paired sample t-tests were performed on predicted and measured values from each model to determine the optimal fitting degree.

1.8 Spatial Trend Distribution Mapping

Using the optimal model, Scr reference values for healthy adults were predicted for 2,322 cities and counties across China. GIS software was then applied for Kriging interpolation [18] to generate a spatial trend distribution map of Scr reference values for healthy adults in China.

Results

2.1 Spatial Autocorrelation Analysis of Scr Reference Values

Using ArcGIS 10.2 software for spatial autocorrelation analysis, the global Moran' s Index yielded a P-value of 0.000. With a standardized statistic Z = 34.3613 at the 0.01 confidence level, the results indicated significant spatial correlation, demonstrating that the distribution of Scr reference values was not random.

2.2 Correlation Analysis Between Scr Reference Values and Geographic Factors

Using SPSS 22.0 software with Scr reference values as the dependent variable and 23 geographic factors as independent variables, correlation analysis revealed single correlation coefficients (r) and P-values between Scr reference values and each geographic factor . The results showed that among the 23 geographic indicators, seven factors correlated significantly with Scr reference values: latitude, annual sunshine duration, annual average temperature, annual average relative humidity, annual precipitation, annual temperature range, and topsoil (silt) cation exchange capacity.

2.3 Principal Component Analysis

2.3.1 Suitability Assessment for Factor Analysis Before conducting principal component analysis, factor analysis suitability was assessed. Using SPSS 22.0, the KMO measure of sampling adequacy for the seven geographic factors was 0.834, indicating suitability for factor analysis as the value approached 1. Additionally, Bartlett' s test of sphericity yielded a significance value of 0.000 (less than 0.05), rejecting the null hypothesis and confirming correlations among variables, thus confirming suitability for factor analysis.

2.3.2 Principal Component Extraction Based on the statistical information of principal components, including initial eigenvalues, contribution rates, and cumulative contribution rates, the first principal component contributed 70.766% of the variance, while the second principal component had an eigenvalue of 0.920 (close to 1) and explained 13.144% of the total variance among the seven original variables. The cumulative contribution rate of the first two eigenvalues was 83.910%; therefore, these two principal components (designated Z1 and Z2) were extracted to replace the seven original indicators.

2.3.3 Principal Component Regression Analysis Using SPSS 22.0 with principal components Z1 and Z2 as independent variables and Scr reference values as the dependent variable, regression analysis yielded the equation:

$$\hat{Y} = 72.00 - 0.02300Z_1 - 0.05300Z_2 \pm 10.49$$

Substituting the expressions for principal components Z1 and Z2 into this regression equation produced the linear regression model for Scr reference values with the seven geographic factors:

$$\hat{Y} = 72.00 - 0.009360X_1 - 0.00003000X_4 + 0.01231X_5 - 0.002690X_6 + 0.003070X_7 - 0.01215X_8 - 0.05403X_{19} \pm 10.49$$

where $\hat{Y}$ represents the Scr reference value for healthy adults and 10.49 is the residual standard deviation.

2.4 Ridge Regression Analysis of Scr Reference Values

Using Scr reference values as the dependent variable and the seven geographic factors as independent variables, ridge regression analysis was performed. A ridge trace plot was constructed with ridge parameter K on the horizontal axis and regression coefficients for each factor on the vertical axis [Figure 2: see original paper]. The plot showed that when K = 0.3, the ridge traces stabilized, so this value was selected as the ridge parameter. Using SAS software for computational analysis, the regression equation between Scr reference values and the seven geographic factors was:

$$\hat{Y} = 76.35 - 0.1122X_1 - 0.0001000X_4 + 0.09462X_5 + 0.02603X_6 + 0.00006000X_7 - 0.03169X_8 - 0.1066X_{19} \pm 10.50$$

where $\hat{Y}$ represents the Scr reference value for healthy adults and 10.50 is the residual standard deviation.

2.5 Selection of Optimal Model

The two models derived from principal component analysis and ridge regression analysis were used to predict Scr reference values for healthy adults across different regions of China. Paired sample t-tests on predicted and measured values were performed using SPSS 22.0. Results showed t = 0.144, P = 0.885 for the principal component model, and t = 0.000, P = 1.000 for the ridge regression model. Since both P-values exceeded 0.05, both models demonstrated good consistency between predicted and measured values. However, the ridge regression model (P = 1.000) showed superior predictive performance compared to the principal component model.

Additionally, data from several provincial capitals were randomly selected to compare measured values with predictions from both models [Figure 3: see original paper]. The ridge regression model predictions were notably closer to measured values than those from the principal component model. Based on these paired t-test results and comparative analysis, the ridge regression model was selected as the optimal predictive model for Scr reference values in healthy adults.

2.6 Geographic Spatial Distribution of Scr Reference Values

Using the optimal model, Scr reference values for healthy adults were predicted for 2,322 cities and counties across China. Kriging interpolation was performed using ArcGIS 10.2 software to generate the spatial trend distribution map [Figure 4: see original paper]. The map illustrates regional variations in predicted Scr reference values and their spatial distribution patterns nationwide. Different color tones represent different Scr value ranges, with similar tones indicating comparable reference values between regions and contrasting tones indicating substantial differences.

Correlation analysis identified seven geographic indicators associated with Scr reference values, among which latitude, annual average temperature, annual average relative humidity, annual precipitation, and annual temperature range showed extremely significant correlations. Following model construction, prediction, and interpolation mapping, the distribution pattern revealed higher Scr reference values in southern China and lower values in northern China, with regular fluctuations corresponding to latitude. Values were relatively low in northeastern and northwestern regions, while southwestern and southeastern coastal areas showed higher values. Latitude, as the most significant correlate, was clearly reflected in the predicted distribution map.

Discussion

Among the seven geographic factors correlated with Scr reference values, most were climate indicators, suggesting that climate is the primary factor influencing Scr reference values in healthy adults. Traditional Chinese medicine theory may explain this climate effect. Southern China is warm, humid, and rainy, where external pathogenic factors predominantly manifest as warm diseases with damp-heat pathologies. Heat injures qi while dampness obstructs the spleen, making spleen deficiency with phlegm-dampness constitution common among southern populations. Damp-heat pathogens tend to invade the middle and lower jiao, potentially causing kidney dysfunction and elevated Scr levels [19]. Northern China has dry, cold climates where external pathogenic factors mainly involve wind-cold, which primarily affects the upper respiratory tract with minimal impact on the urinary system. Southern China’s subtropical and tropical monsoon climate, characterized by high annual average temperature, humidity, and precipitation, corresponds to higher Scr reference values, consistent with the color patterns in [Figure 4: see original paper].

Latitude primarily influences Scr reference values through its effect on temperature conditions, with temperatures decreasing as latitude increases from south to north in China [20]. Combined with the aforementioned traditional Chinese medicine theory, this explains why Scr reference values are higher in southern, especially southeastern, regions compared to northern areas, aligning with the color variation patterns in [Figure 4: see original paper].

Soil environmental effects on Scr reference values are not direct. Among soil indicators, only topsoil (silt) cation exchange capacity correlated with Scr reference values. This parameter is an important indicator of soil fertility and primarily affects human health through the food chain, indirectly reflecting how dietary differences influence normal physiological values [21-22].

Geographic environments constitute the habitat for human survival. Different regions have distinct climates and soil conditions, leading to variations in human constitution, pathogenic factors, and disease mechanisms, which in turn affect human physiology. This results in obvious regional differences in many physiological characteristics, including kidney function indicators [23-25]. The traditional Chinese medicine concept of “three-factor appropriateness” (treatment tailored to season, locality, and individual) posits that disease occurrence, development, and prognosis are influenced by multiple factors including seasonal climate and geographic environment. Therefore, to achieve expected therapeutic outcomes, treatment must be formulated according to the relationships among disease, climate, geography, and patient [26-27]. This study’ s development of predictive models for regional Scr reference values and selection of an optimal model to estimate values for 2,322 cities and counties aligns with this “three-factor appropriateness” philosophy.

Current medical reference value standards do not consider geographic environmental influences, which may compromise their accuracy and scientific validity. Incorporating geographic factors into medical analysis enables region-specific determination of reference value ranges for healthy populations, profoundly improving clinical diagnostic accuracy. However, the mechanisms through which geographic factors influence Scr reference values require further investigation.

References

[1] Jin YQ. Analysis of clinical value of serum cystatin C and serum creatinine detection in diagnosing kidney injury [J]. Contemporary Medical Symposium, 2015, 13(24): 44-5.

[2] Wang GP, Li H, Niu H, et al. Value of serum cystatin C combined with serum creatinine in diagnosing hepatorenal syndrome [J]. Chinese Journal of Current Advances in General Surgery, 2015, 18(3): 247-8, 252.

[3] Pavkov ME, Knowler WC, Hanson RL, et al. Comparison of serum cystatin C, serum creatinine, measured GFR, and estimated GFR to assess the risk of kidney failure in American Indians with diabetic nephropathy [J]. Am J Kidney Dis, 2013, 62(1): 33-41.

[4] Yan CN, Yan JH, Song JL, et al. Dictionary of Chinese Cities and Counties [M]. Beijing: Central Party School Press, 1991.

[5] Yan H, Shen GQ, Mao YS, et al. Climate Atlas of the People' s Republic of China [M]. Beijing: Meteorological Press, 2002.

[6] Luo JH. Clinical significance of serum parathyroid hormone and bone mineral density detection in diabetic nephropathy patients [J]. Journal of Ningxia Medical University, 2015, 37(9): 1074-7.

[7] Song Y. Application value of serum cystatin C in early diagnosis of diabetic nephropathy [J]. Jilin Medical Journal, 2015, 36(2): 285-285.

[8] Zhang JB, Ma LT. Value of combined detection of serum cystatin C and $ $2-microglobulin in early diagnosis of diabetic nephropathy [J]. Shaanxi Medical Journal, 2015, 44(2): 240-1.

[9] Bo T, Huang H, Zhang BZ. Correlation analysis of serum Cys-C, homocysteine and diabetic nephropathy [J]. China Modern Medicine, 2015, 25(15): 104-5, 108.

[10] Chen HX, Wang TB, Yang Q, et al. Relationship between serum cystatin C and blood pressure: a report of 912 cases [J]. Journal of Southern Medical University, 2015, 35(7): 1055-8.

[11] Wu LT, Liu MQ. Study on cystatin C combined with eGFR-MDRD in diagnosing early DN [J]. Chinese Journal of Immunology, 2015, 31(10): 1394-6, 1399.

[12] Ding LH. Clinical observation and study of serum Cys-C, $ $2-MG, BUN, and SCr in early diabetic nephropathy patients [J]. Journal of Shandong Medical College, 2014, 36(1): 54-6.

[13] Ge M. Research progress on the relationship between normal medical reference values and geographic factors [J]. Foreign Medical Sciences: Section of Medical Geography, 2009, 30(2): 49-52, 92.

[14] Han X. Relationship between mitral flow spectrum E/A reference values of Chinese adults and geographic environment [D]. Xi' an: Shaanxi Normal University, 2015.

[15] Miao G. Normal reference value of hemoglobin of older boys and geographical factors [J]. Biosci Rep, 2004, 23(5/6): 305-12.

[16] Zeng Q, Zhou YR. Ridge regression analysis and its application in analyzing the impact of atmospheric environmental factors on population health [J]. Chinese Journal of Health Statistics, 1997(4): 25-7.

[17] Sheffield L, Whitford H, Cressie NA. Use of paired sample T-Test in the real world [J]. Practotol, 2009, 102(3): 201-5.

[18] Tang GA, Yang X. ArcGIS Geographic Information System Spatial Analysis Experimental Tutorial [M]. Beijing: Science Press, 2006: 50-203.

[19] You J. Study on screening of biological markers for evaluating stress load of military operations in hot environments and prevention of damage [D]. Chongqing: Third Military Medical University, 2012.

[20] Liu DY. Discussion on physiological characteristics and adaptability of human body in Qinghai-Tibet Plateau region [J]. Journal of Qinghai Normal University: Natural Science Edition, 1980(1): 52.

[21] Ge M, Xue RY, He JW, et al. Geographic distribution of peak expiratory flow reference values for Chinese male children [J]. Geographical Research, 2014, 33(3): 451-66.

[22] Wang YG. Soil environment and human health [J]. Inner Mongolia Petrochemical Industry, 2009(24): 88-9.

[23] Miao G, Zhang Y, He J, et al. Normal reference value of red blood cell count of Chinese young men and geographical factors [J]. Semin Diagn Pathol, 2009, 26(1): 53-60.

[24] Yang LS, Wang WY, Tan JA, et al. Research achievements and prospects of environmental geography and human health [J]. Geographical Research, 2010, 29(9): 1571-83.

[25] Liu XL, Ge M, Wang CX, et al. Analysis of the influence of geographic environment on P wave dispersion reference values [J]. Journal of Northeast Normal University: Natural Science Edition, 2015, 47(1): 135-40.

[26] Wang XQ, Cheng XK. Discussion on the thought of "three-factor appropriateness" in *Neijing* [J]. Journal of Yunnan University of Traditional Chinese Medicine, 2010, 33(3): 16-7.

[27] Li ZG, Jin XL. Examples of "three-factor appropriateness" thought in the four great physicians of the Jin and Yuan Dynasties [J]. Beijing Journal of Traditional Chinese Medicine, 2009, 28(1): 25-7.

(Edited by Wu Jinya)

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.