Grey Correlation Analysis Between Liver Cancer and Drinking Water Quality in Southeastern Coastal Areas of China

NI Fu-quan[1], XU Li-ping[1] , FU Cheng-wei[1], GUO Shu-long[2]

[1]College of Information & Engineering,
Sichuan Agricultural Univ.,
Ya'an, China
nfq1965@163.com

[2]China Water Northeastern
Investigation Design &Research Co.Ltd.
Changchun,China

***Abstract*—Using grey correlation analysis to the relationship between the incidence of liver cancer and drinking water quality in Fujian , Jiangsu , and Guangxi provinces, considering the basic features of the original and the secondary environments in the study areas, the following conclusions can be drawn: liver cancer can be caused by many factors and drinking water may be among the most important ones; minerals (e.g. Ca^{2+}, Mg^{2+}, Mo, Zn, etc.) needed in drinking water by human body in southeastern coastal areas of China are relatively low; Ca^{2+} , Cu ,NO_2^- etc . are closely related to the incidence of liver cancer.**

Keywords-Grey Correlation Analysis; Incidence of Liver Cancer; Drinking Water Quality; Southeastern Coastal Areas of China

I. INTRODUCTION

In order to research the key factors that control the drinking water quality in the disease district, conventional methods, such as step by step regression analysis, principal component analysis, etc. are usually adopted to study the relationship between drinking water quality and the incidence of the endemic liver cancer. However, drinking water quality system has uncertainty as well as the relationships among factors. As a result, the system is a grey system and the grey system theory [1] should be used to deal with its relationship.

II. THEORY AND CALCULATION

Grey correlation analysis is used to study the relationship among factors based on similarity or difference. Usually the grey correlativity is used to express the affected degrees among factors. The larger the correlativity is, the closer relation and influence between that factor and others.

The following is an example of this procedure.

Supposing an original data matrix of impact factors as follows:

$$X_{ij} = \begin{vmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{vmatrix} \quad (1)$$

Where, X_{ij} is the value of index j about sample i, n is the number of samples, m is the number of indices.

Standardizing the original data matrix:

$$X\ i(k)\ =\ \frac{X\ i(k)}{X\ i(1)} \qquad (2)$$

Where, $i = 1,2,...,n$; $k = 1,2,...m$.

Calculating the absolute value of difference between mother-sequence and son-sequence:

$$d_i(k) = |X_0(k) - X_i(k)| \qquad (3)$$

Where, $X_0(k)$ and $X_i(k)$ are selected mother-sequence and son-sequence respectively; $d_i(k)$ is the absolute value of difference between mother-sequence and son-sequence.

Calculating the largest and smallest difference between the two sequences:

$$\Delta\min = \min(i)\min(k)(d_i(k))$$
$$\Delta\max = \max(i)\max(k)(d_j(k)) \quad (4)$$

Where, $\Delta\min$ is the smallest data among the smaller datum; $\Delta\max$ is the largest data among the larger datum. Namely, in the case of i, the smallest (or the largest) value among $d_i(k)$ datum is picked out, and then the smallest (or the largest) multiplied value among them is obtained.

Calculating the correlation coefficient:

$$L_i(k) = \frac{\Delta\min + C \times \Delta\max}{d_i(k) + C \times \Delta\max} \quad (5)$$

Where, C is a constant (C is 0.5 in this paper), $L_i(k)$ is the correlation coefficient. Other symbols have been given previously.

Calculate the correlation degree:

$$R_i = \frac{1}{m}\sum_{k=1}^{m} L_i(k) \quad (6)$$

Where, R_i is the correlation degree. Other symbols have been given previously.

III. CASE

Liver cancer is unevenly distributed in China. High-risk Area for it is the southeastern coastal areas, such as Fusu county (Guangxi province), Qidong county (Jiangsu province), Daishan county (Zhejiang province), Shanghai etc. The mortality rates of liver cancer respectively are 48.83×10^{-5},47.93×10^{-5} ,36.5×10^{-5} and 21.17×10^{-5} [2,3]. With respect to the cause of liver cancer, the "water and soil" doctrine holds: concentration of microelements (e.g. Mo,Se,Cu,Zn etc.) of water and soil in affected area are relatively low; however, concentration of NO_2^-, NH_4^+ are relatively high. In addition, from the point of biological function of elements , Ca^{2+}, Mg^{2+}, Mo, Se, Cu, Ge, Zn etc. are closely related to liver cancer .The pollution of drinking water shouldn't be neglected , the increasing concentration of NO_2^-, NH_4^+, the phenol, COD and humic-acid. Because of this background, Ca^{2+}, Mg^{2+}, Mo, Ba, Cu, Zn, NO^-_2 , NH_4^+ , COD, and humic-acid in drinking water are selected as indices to conduct grey correlation analysis to liver cancer.

A.The Original Data

The tested indices of water quality in Changping country, Quli country,Youth Pond in Fusui county of Guangxi province are Ca^{2+}, Mg^{2+}, Mo, Cu , Zn.

The tested indices of surface water (pond water, irrigation water and river water) and ground water (shallow well water, deep well water) in Qidong county of Jiangsu province are NO_2^- , COD and humic-acid.

The tested indices of groundwater quality in Changle county of Fujian province are Cu, Ba, NO_2^-, NH_4^+ and COD etc.

B.Results

The results of the grey correlation analysis are given in table 1.

C. Discussion and Conclusions

From table 1, the following conclusions are drawn:

Both micro-elements and macro-elements of drinking water in Fusui county of Guangxi province are closely related to the incidence of liver cancer and the correlation degrees of Ca^{2+}, Mg^{2+}, Mo, Zn are 0.708, 0.523, 0.556 and 0.521 respectively.

Pollution indices, such as humic-acid, COD,NO_2^- in drinking water in Qidong county of Jiangsu province are obviously related to the incidence of liver cancer and the correlation degrees are 0.777, 0.680 and 0.811 respectively.

The tested indices of micro-elements, such as Cu, Ba in drinking water of the four counties in Fujian province are obviously related to the incidence of liver cancer and the correlation degrees are 0.880,0.624 respectively; pollution indices, such as COD, NH_4^+, NO_2^- are closely

TABLE 1.CORRELATION DEGREES

Tested indices / Tested sites	Ca^{2}	Mg^{2+}	Zn	Cu	Mo	Ba	NO_2^-	COD	NH_4^+	Humic acid
Fusui,Guangxi	0.708	0.523	0.521	0.377	0.566					
Qidong,JiangSu							0.811	0.680		0.777
Changle etc. ,in Fujian				0.880		0.624	0.747	0.767	0.791	

related to the incidence of liver cancer and the correlation degrees are 0.791,0.767,0.747 respectively. The micro-element Cu is very special; its correlation degree (0.880) is the largest in Fujjian province rather than the smallest value 0.377 in all correlation degrees in Guangxi province. The incidence of liver cancer is affected by different percentages of mineral elements in drinking water.

All results above show that different drinking water in different areas has different influence on the endemic liver cancer. So the relationship should be analyzed with condition.

From the point of environmental and hydrogeological characteristics, outcropping strata in Fusui county of Guangxi province are as follows: grey white thick strata, massive limestone accompanied by dolostone, colitic limestone and bioclastic limestone in Changping; grey-white thick-bedded massive limestone, dolomitic limestone in Quli. Because of water shortage, local inhabitants mainly drink low-quality pool water or well water leaking from the pool; moreover, the climate type is the moist-hot, limestone mother material rock weathers into red soil easily which contains a few carbonate. As a result, this environment is lack of various micro-elements and macro-elements, such as Ca^{2+},Mg^{2+},Zn ,Cu ,Mo, etc. See in table 2.

TABLE 2.CHEMICAL COMPOSITION OF LOWER CARBONIFEROUS MOTHER SOIL

Composition	SiO_2	Al_2O_3	Fe_2O_3	FeO	CaO	MgO	K_2O	Na_2O	TiO_2	MnO
Concentration (%)	63.2 | 79.8	7.4 | 15.3	4.7 | 6.4	0.05 | 0.82	0.1 | 1.18	0.42 | 1.82	2.57 | 3.31	0.05 | 0.10	0.38 | 0.68	0.006 | 0.25

Qidong county is a conjunction of alluvion, silting, littoral deposit and continental-oceanic interacting strata. Concentration of humus in soil is higher, belonging to reduction environment. There is methane (mainly CH_4) in swampland and the margin of ponds. Pollution material of industry and living added, the values of NO_2^- , NH_4^+, COD in drinking water are relatively higher [4].

For the four counties of Fijian province, harmful and noxious material are rich in the ground water, ditch water, underground water of shallow loose rock type, especially in the ground water of Quaternary period marine-deposition strata. Because of the flattish topography, the groundwater flow is very slow. Furthermore, the flow ability and self-purification of the pollution material are very limited. Bedrock of high risk area of liver cancer is granite and ash tuff. Concentration of Barium in water of this area is higher.

To sum up, the high risk area of liver cancer is based on the forming of the original environment and the affection of the secondary environment, in which, the former includes the condition of geomorphology, geology and geochemistry and the latter is mainly showed in transformation and utilization to environment by human beings. The liver cancer area is the result of many factors that interact with each other. High or low of concentration of Ca^{2+}, Mg^{2+}, Zn, Cu, Mo, Ba in drinking water can affect incidence of liver cancer and induce different cancer and tumor. Meanwhile, pollution of drinking water shouldn't be neglected, NO_2^-, NH_4^+ in particular. Experiments show that NO_2^- combining with the phenol evolved easily into nitroso compound. When drinking water contains high organic material and the COD value is high, big change of NO_2^- plays an active role in the formation of nitroso compound as well as the inducing and growing of cancer cell in human body. Consequently, high NO_2^-, NH_4^+, COD, humic-acid in drinking water may be the common and important causes of liver cancer .

The cause of endemic liver cancer is quite complicated; no unanimous conclusions can be drawn. The following several aspects are worth further discussing：

Endemic liver cancer is not absolutely controlled by only one factor in all affecting factors that interact with each other. Endemic liver cancer is closely related to chemical compositions in drinking water. It is worth further studying.

The higher concentration of NO_2^-, NH_4^+, COD and humic-acid in drinking water can lead to higher incidence of liver cancer. High or low concentration of Ca^{2+}, Mg^{2+}, Zn, Cu, Mo, Ba can affect incidence of liver cancer and induce different cancer and tumor. As a kind of new method, although the grey correlation analysis, quantifies the relationship between incidence of liver cancer and concentrations of chemical compositions, it still belongs to the qualitative analysis, i.e. it is a quantized qualitative analysis, which providing a kind of system analysis for analyzing and judging the impact factors of the endemic liver cancer.

Endemic liver cancer, which is related to many factors, such as local living conditions, living habits, even inherits, sex, age, occupation etc. needs further research [5,6].

REFERENCES

[1] Deng Julong, 1983, *Grey System Theory,* The middle China Physics Industry University Press, Wuhan, China .48

[2] Wang Xuquan, V., 1992, *Study on the relation-ship between carcinoma of esophagus and drinking water quality in the four counties, such as Changle, ect. China,* Environment Science. v. 12, 144-47.

[3] Nianfeng Lin, 1991, *Medical Environmental Geochemistry*. Jilin Science and Technology Press,ChangChun,China. 395

[4]Chun-Yuh Yang, Deng-Chuang Wu, Chih-Ching Chang.*Nitrate in drinking water and risk of death from colon cancer in Taiwan*. Environment International 33 (2007) 649-653.

[5] YU Shunzhang，M U Lina，CAI Ling，etc.The drinking water and three environm ental risk factors for hepatOcellular carcinoma in Taixing by case-control study. Fudan Univ. J. Med. Sci.,2008,JAN.35(1),31-38.

[6]John-Mark Davies, Asit Mazumder. Health and environmental policy issues in Canada: the role of watershed management in sustaining clean drinking water quality at surface sources. Journal of Environmental Management 68 (2003) 273–286.