PO Box 2345, Beijing 100023, China
Fax: +86-10-85381893
E-mail: wjg@wjgnet.com www.wjgnet.com

ELSEVIER

World J Gastroenterol 2005;11(5):641-644
World Journal of Gastroenterology ISSN 1007-9327

Quantitative assessment model for gastric cancer screening

Kun Chen, Wei-Ping Yu, Liang Song, Yi-Min Zhu

Kun Chen, Wei-Ping Yu, Liang Song, Yi-Min Zhu, School of Medicine, Zhejiang University, Hangzhou 310031, Zhejiang Province, China
Supported by National Nature Science Foundation of China, No. 30170828
Correspondence to: Kun Chen, M.D., Department of Epidemiology, School of Medicine, Zhejiang University, 353 Yan-an Road, Hangzhou 310031, Zhejiang Province, China. ck@zjuem.zju.edu.cn
Telephone: +86-571-87217190 **Fax:** +86-571-87217184
Received: 2004-03-03 **Accepted:** 2004-04-13

Abstract

AIM: To set up a mathematic model for gastric cancer screening and to evaluate its function in mass screening for gastric cancer.

METHODS: A case control study was carried on in 66 patients and 198 normal people, then the risk and protective factors of gastric cancer were determined, including heavy manual work, foods such as small yellow-fin tuna, dried small shrimps, squills, crabs, mothers suffering from gastric diseases, spouse alive, use of refrigerators and hot food, *etc*. According to some principles and methods of probability and fuzzy mathematics, a quantitative assessment model was established as follows: first, we selected some factors significant in statistics, and calculated weight coefficient for each one by two different methods; second, population space was divided into gastric cancer fuzzy subset and non gastric cancer fuzzy subset, then a mathematic model for each subset was established, we got a mathematic expression of attribute degree (AD).

RESULTS: Based on the data of 63 patients and 693 normal people, AD of each subject was calculated. Considering the sensitivity and specificity, the thresholds of AD values calculated were configured with 0.20 and 0.17, respectively. According to these thresholds, the sensitivity and specificity of the quantitative model were about 69% and 63%. Moreover, statistical test showed that the identification outcomes of these two different calculation methods were identical (P>0.05).

CONCLUSION: The validity of this method is satisfactory. It is convenient, feasible, economic and can be used to determine individual and population risks of gastric cancer.

Key words: Gastric cancer; Mass screening; Quantitative assessment model

Chen K, Yu WP, Song L, Zhu YM. Quantitative assessment model for gastric cancer screening. *World J Gastroenterol* 2005; 11(5): 641-644
http://www.wjgnet.com/1007-9327/11/641.asp

INTRODUCTION

Gastric cancer is the second most common cause of cancer deaths in the world, and China is one of the high-risk areas[1]. Population screening is an effective program for providing early diagnosis and subsequent treatment of gastric cancer at its curable stage. Whatever screening method is used, the most important thing is that the method should be convenient, feasible, and acceptable to the target population[2]. At present, the methods used to find and diagnose gastric cancer at early time are complicated, or their sensitivity and specificity are dissatisfactory. Based on a population case-control study, a mathematic model was established for determining individual and population risks of gastric cancer in this paper. An assessment of its practical application was also carried out to determine its validity.

MATERIALS AND METHODS

Case-control study

To study the risk factors for gastric cancer, a case-control study including 66 patients and 198 normal people, was carried out in 1999. Factors involving demographic variables, diet, drinking water source, individual habits, disease history and family history of gastric cancer were investigated in this study. Risk and protective factors for gastric cancer were determined by the fast epidemiology assessment method[3]. At the level of $\alpha = 0.10$, gastric cancer risk factors included heavy manual work (>2 h/d), foods such as small yellow-fin tuna, dried small shrimps, squills, crabs, and mothers suffering from gastric diseases. In contrast, spouse alive, use of refrigerators and hot food were the protective factors against gastric cancer.

Mathematic expression

Based on the case-control study, a quantitative assessment method was put forward by selecting some factors significant in statistics, including risk factors and protective factors, with application of some theories and approaches of fuzzy and probability mathematics. The method was set up as follows. Population characteristic space was divided into gastric cancer fuzzy subset and non gastric cancer fuzzy subset, respectively. Which subset each subject belonged to was determined by attribute function, and the determination probability should be maximal, or its error probability should be minimum. $\underset{\sim}{A}$ was configured as a fuzzy subset suffering from gastric cancer. First, for setting up a fuzzy mathematical model, a group of standard factors should be determined, that was U_i. Weight sum (P) of U_i was configured as following:

$$P\underset{\sim}{\overset{\circ}{A}}=\sum_{j=1}^{n}\alpha_j C_j \cdots\cdots (1)$$

In the expression (1), j = 1, 2, 3,...n, and n is the number of the factors selected. C_j is an identification score of each factor, that is C_j equals 1 when a subject has a factor of F_i, no matter that F_i is a risk factor or not. α_j is an attribute coefficient of F_i, thereof P is weight sum of all factors (F_i).

$P\underset{\sim}{A}$ is representative of weight sum of a subject when he (or she) has some or all factors, that is:

$$P\underset{\sim}{A}=\sum_{j=1}^{n}\alpha_j C'_j \cdots\cdots (2)\ (j = 1, 2, 3\ldots, n)$$

In the expression (2), α_j is the attribute coefficient of each

F_i, and when a subject has a certain F_i, C'_j equals C_j if F_i is a risk factor. On the contrary, C_j equals 0 or 1 if a subject has no F_i. An attribute function (individual or population) can be set up with $P\tilde{\tilde{A}}$ and $P\tilde{\tilde{A}}$:

$$\mu_{\tilde{\tilde{A}}}(U_i) = \frac{P_{\tilde{\tilde{A}}}}{P_{\tilde{\tilde{A}}}} \cdots\cdots \quad (3)$$

In the expression (3), $\mu_{\tilde{\tilde{A}}}(U_i)$ is the expression of attribute function of $\tilde{\tilde{A}}$. A specific value ranging from 0 to 1 can be calculated by the expression for each subject, that is AD.

Determination of weight coefficient for each factor

α_j in the expressions (1) and (2) was calculated by two different methods. In method 1, OR for each F_i is taken as the weight coefficient of each F_i (if F_i is a protective factor, its weight coefficient equals 1/OR), that is $\alpha_i = \frac{OR_i (or 1/OR_i)}{\sum OR_i}$. A weight coefficient was calculated using the conditional probability and entropy of each F_i, which is $\alpha_i = \frac{H(F_I / D)}{H(F_I / D + H(F_I / \overline{D})}$ in method 2. In the expression, $H(F_i/D)$ is the entropy of F_i in the condition of patients, and $H\left(F_i / \overline{D}\right)$ is entropy of F_i in the condition of normal people.

$$H(Fi / D) = -P(Fi / D)\log_2 P(Fi / D)$$

$$H(Fi / \overline{D}) = -P(Fi / \overline{D})\log_2 P(Fi / \overline{D})$$

Weight coefficients of all factors are illustrated in Table 1.

Assessment of quantitative method

AD values of 63 patients and 693 normal people were calculated with the expression (3), based on the quantitative assessment of individual risk and population screening of gastric cancer. The variational trend of sensitivity, specificity and Youden indexes were identical, which are illustrated in Figure 1, though calculation methods were different.

Because AD is a continuous variable, the identification threshold could be determined based on actual needs. The threshold could be reduced a little in order to increase the positive rate while trying to check out more patients. Moreover, the threshold could be raised to increase the specificity and reduce the rate of false diagnosis in detective diagnosis (Figure 1).

In order to get the maximal Youden indexes, the thresholds of AD values were configured with 0.20 and 0.17, considering the sensitivity and specificity of population screening for gastric cancer. Diagnostic value of different calculation methods of weight coefficients are summaried in Table 2, and significance test showed that these Youden indexes had no statistical significance (P>0.05). Thus we could see the outcomes tended to be identical.

DISCUSSION

Gastric cancer, the most common fatal malignancy in the world, causes more than 750 000 deaths annually[4]. To the year of 2005, the mortality of gastric cancer is about to reach 26.3/100 000 per year in china[5]. Early finding, diagnosis and treatment are the keys to reduce the mortality of gastric cancer, to raise the survival rate and improve the life quality of patients.

Table 1 OR values, confidence limits and weight coefficients of risk factors and protective factors of gastric cancer

Variable	OR_i	95%C.I.	α_i (method 1)	α_i (method 2)
Time of heavy manual work (>2 h/d)	2.00	1.14-3.54	0.084	0.439
Eating small yellow-fin tuna frequently	1.52	1.10-2.09	0.064	0.490
Often eating squills (dry)	6.12	1.15-32.66	0.257	0.782
Eating dried small shrimps frequently	1.26	0.99-1.60	0.053	0.462
Often eating squills (fresh)	1.70	0.92-3.13	0.071	0.429
Eating crabs frequently	1.76	1.00-3.09	0.074	0.451
Mother suffering tummy bug	5.51	0.95-32.06	0.231	0.774
Eating shortly after anger	2.07	1.31-3.27	0.087	0.522
Spouse alive	0.89	0.81-0.99	0.047	0.819
Using refrigerators	0.47	0.19-1.15	0.089	0.399
Often eating hot food	0.52	0.28-0.97	0.081	0.438

Table 2 Diagnosis value of different calculation method of weight coefficient

αj	AD	Patients (n/N)	Normal people (n/N)	Sensitivity (%)	Specificity (%)	Youden index
Method 1	≥0.20	44/64	440/693	69.8	63.5	0.333
Method 2	≥0.17	43/63	435/693	68.3	62.8	0.311

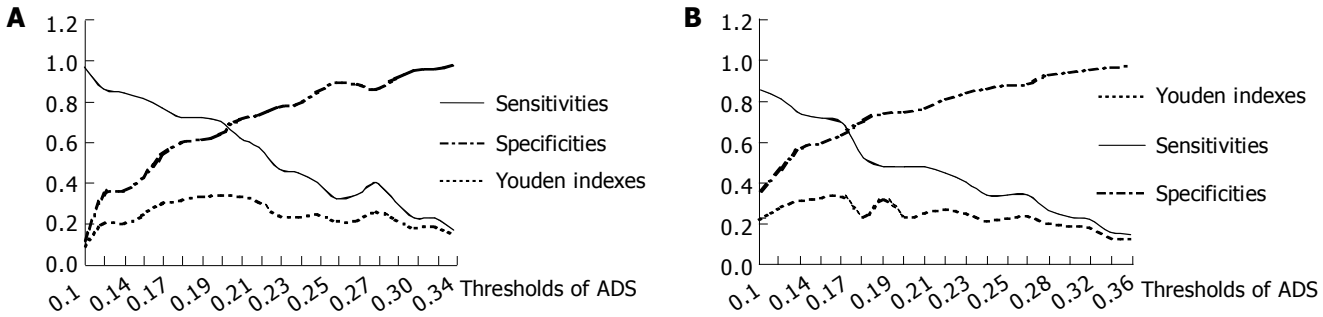

Figure 1 Diagnostic value of risk assessment for gastric cancer by methed 1 (A) and method 2 (B).

In early gastric cancer, the 5-year survival rate is greater than 90% if treated by experienced hands[6]. It has important significance to develop a simple and feasible screening method to find high risk populations or early gastric cancer.

Up to now various methods for gastric cancer screening have been developed. X-ray examination[7,8], endoscopy[9,10], Helicobacter pylori screening[11], power Dopller imaging[12] and photo fluorography[13] are used in gastric cancer screening. In addition, gastric occult blood bead test[14], serum pepsinogen concentration measure[15] and fecal carcinoembryonic antigen measurement[16], are also used in screening gastric cancer. The efficacy of some methods has reached an ideal level[15-17]. Except for the methods mentioned above, some gene alterations, such as hMLH1 methylation[18], BAT-26 mutation localized in intron 5 of hMSH2 gene[19], E-cadherin germline mutations[20], cyclin overexpression, microsatellite instability, P53 mutation[21] are useful molecular markers for gastric cancer.

Quantitative method has rarely been applied to gastric cancer screening either at home or at abroad. Qiu *et al*[22] studied the application of pattern recognition method in 1994. A computer program was designed according to the principle of pattern recognition and risk factors for gastric cancer. Its detection rate was 1.54/1 000 in a study of 51 735 males aged 45-64 years.

Though the accuracy of some screening methods is ideal, they have obvious disadvantages in practice. Endoscopy and biopsy could make subjects discomfort. Many suspicious patients tend to refuse these kinds of examinations. Another shortcoming is the high cost. Some molecular biology marker tests cost subjects too much due to expensive reagents, some of them are invasive because gastric juice must be collected before tests. In addition, the pattern recognition method reported by Qiu *et al*[22] was complicated and high-cost, because 61 indexes must be questioned to subjects. Moreover, the principle of pattern recognition is difficult to be mastered by subjects and inquirers, and the workload is heavy. Therefore, a conclusion may be reached that these methods are not suitable for application in China.

Compared with these methods, the quantitative assessment model is simple, economic, non-invasive and feasible. Identification can be run as long as each subject fills in a simple questionnaire, and calculation method is simpler than traditional mathematic methods such as regression identification. Furthermore, the diagnostic value of this quantitative method is relatively high. Its sensitivity and specificity are about 69% and 63%. Given the factors just outlined, this quantitative screening method for gastric cancer can be applied to a large population in China.

Formerly, quantitative methods were mostly applied to differential diagnosis in clinic[23]. There have been some methods for assessing health hazard/health risk based on epidemic study of population since 1980s[24]. In the 1990s, a quantitatively scored cancer-risk assessment model[25] was developed to promote cancer prevention and screening in America. Subsequently quantitative models are widely used in the diagnosis, treatment and prevention of diseases, such as multi-stage lung cancer[26], assessment of cancer death in elderly patients[27], quantitative model for early diagnosis of colorectal cancer[28] and efficacy evaluation of intervention experiments[29]. In the late 1980s, Chen *et al*[30] reported a mathematic model for mass screening of colorectal cancer, which was subsequently proved to be a convenient, effective and economic screening method. In recent years, the mathematic model has been applied to screening other diseases, including coronary heart disease and stroke[31], lung cancer[32]. These studies show that the model has good efficacy. However, there are two points to which attention must be paid. One is how to identify threshold values. In our study, the threshold could be defined with practical application because AD is a continuous variable. However, further follow-up is needed to increase its precision in screening other diseases, because causes of different diseases are complicated. The other is the low Youden index[30,31]. The reasons why many factors are associated with these diseases are still unclear.

In short, the quantitative method can be regarded as the front line method for assessing risks of gastric cancer. Occurrence of gastric cancer is the outcome of many influential factors, which are possibly different in different areas and populations, that combining with actual status is very important. Further studies are needed to testify whether this screening method can contribute to the decrease of gastric cancer mortality.

REFERENCES

1 **Correa P,** Piazuelo MB, Camargo MC. The future of gastric cancer prevention. *Gastric Cancer* 2004; **7**: 9-16
2 **Tan ZH.** Modern epidemiology. 1st ed. *Beijing: People's Medical Publishing House* 2001: 162
3 **Zhu YM,** Chen K, Zhang Y, Zhu YP, Liu XY, Chen XX, Xu ZZ, Chen JH. Analysis on risk factors of gastric cancer in island residents. *Zhongguo Gonggong Weisheng* 2000; **16**: 916
4 **Xia HH,** Wong BC, Lam SK. Chemoprevention of gastric cancer: current status. *Chin Med J* 2003; **116**: 5-10
5 **Sun XD,** Mu R, Zhou YS, Dai XD, Zhang SW, XM, Sun J, Li LD, Lu FZ, Qiao YL. Analysis of mortality rate of stomach cancer and its trend in twenty years in China. *Zhonghua Zhongliu Zazhi* 2004; **26**: 4-9
6 **Hohenberger P,** Gretschel S. Gastric cancer. *Lancet* 2003; **362**: 305-315
7 **Inaba S,** Hirayama H, Nagata C, Kurisu Y, Takatsuka N, Kawakami N, Shimizu H. Evaluation of a screening program on reduction of gastric cancer mortality in Japan: preliminary results from a cohort study. *Prev Med* 1999; **29**: 102-106
8 **Sasagawa Y,** Sasagawa T, Takasaki K. Mass screening for gastric cancer performed in Costa Rica. *Shokakibyo Gakkai Zasshi* 2002; **99**: 577-583
9 **Han JY,** Son H, Lee WC, Choi BG. The correlation between gastric cancer screening method and the clinicopathologic features of gastric cancer. *Med Oncol* 2003; **20**: 265-269
10 **Riecken B,** Pfeiffer R, Ma JL, Jin ML, Li JY, Liu WD, Zhang L, Chang YS, Gail MH, You WC. No impact of repeated endoscopic screens on gastric cancer mortality in a prospectively followed Chinese population at high risk. *Prev Med* 2002; **34**: 22-28
11 **Roderick P,** Davies R, Raftery J, Crabbe D, Pearce R, Patel P, Bhandari P. Cost-effectiveness of population screening for *Helicobacter pylori* in preventing gastric cancer and peptic ulcer disease, using simulation. *J Med Screen* 2003; **10**: 148-156
12 **Kawasaki T,** Ueo T, Itani T, Shibatohge M, Mimura J, Komori H, Todo A, Kudo M. Vascularity of advanced gastric carcinoma: evaluation by using power Doppler imaging. *J Gastroenterol Hepatol* 2001; **16**: 149-153
13 **Portnoi LM,** Kazantseva IA, Isakov VA, Nefedova VI, Gaganov LE. Gastric cancer screening in selected population of Moscow region: retrospective evaluation. *Eur Radiol* 1999; **9**: 701-705
14 **Zhou L,** Yu H, Zheng S. The value of "occult blood bead" in detection of upper digestive tract disorders with bleeding. *Zhonghua Zhongliu Zazhi* 1999; **21**: 48-50
15 **Kitahara F,** Kobayashi K, Sato T, Kojima Y, Araki T, Fujino MA. Accuracy of screening for gastric cancer using serum pepsinogen concentrations. *Gut* 1999; **44**: 693-697
16 **Kim Y,** Lee S, Park S, Jeon H, Lee W, Kim JK, Cho M, Kim M, Lim J, Kang CS, Han K. Gastrointestinal tract cancer screening using fecal carcinoembryonic antigen. *Ann Clin Lab Sci* 2003; **33**: 32-38
17 **Kubota H,** Kotoh T, Masunaga R, Dhar DK, Shibakita M, Tachibana M, Kohno H, Nagasue N. Impact of screening survey of gastric cancer on clinicopathological features and survival: retrospective study at a single institution. *Surgery* 2000; **128**: 41-47
18 **Waki T,** Tamura G, Tsuchiya T, Sato K, Nishizuka S, Motoyama

T. Promoter methylation status of E-cadherin, hMLH1, and p16 genes in nonneoplastic gastric epithelia. *Am J Pathol* 2002; **161**: 399-403

19 **Wu MS,** Lee CW, Sheu JC, Shun CT, Wang HP, Hong RL, Lee WJ, Lin JT. Alterations of BAT-26 identify a subset of gastric cancer with distinct clinicopathologic features and better postoperative prognosis. *Hepatogastroenterology* 2002; **49**: 285-289

20 **Graziano F,** Ruzzo AM, Bearzi I, Testa E, Lai V, Magnani M. Screening E-cadherin germline mutations in Italian patients with familial diffuse gastric cancer: an analysis in the District of Urbino, Region Marche, Central Italy. *Tumori* 2003; **89**: 255-258

21 **Lam SK.** 9th Seah Cheng Siang Memorial Lecture: gastric cancer--where are we now? *Ann Acad Med Singapore* 1999; **28**: 881-889

22 **Qiu XY,** Shi QX, Shi R, Tu JT, Chen HQ. Mass screening of gastric cancer-establishment of pattern recognition method and its application. *Zhongguo Weisheng Tongji* 1994; **11**: 47-50

23 **Zhou HW,** Sun WM. Metrological medicine of clinic. 1st ed. *Shanghai: Publishing House of Shanghai Medical University* 1999: 51-72

24 **Wagner EH,** Beery WL, Schoenbach VJ, Graham RM. An assessment of health hazard/health risk appraisal. *Am J Public Health* 1982; **72**: 347-352

25 **Lippman SM,** Bassford TL, Meyskens FL. A quantitatively scored cancer-risk assessment tool: its development and use. *J Cancer Educ* 1992; **7**: 15-36

26 **Dawson SV,** Alexeeff GV. Multi-stage model estimates of lung cancer risk from exposure to diesel exhaust, based on a U.S. railroad worker cohort. *Risk Anal* 2001; **21**: 1-18

27 **Walter LC,** Covinsky KE. Cancer screening in elderly patients: a framework for individualized decision making. *JAMA* 2001; **285**: 2750-2756

28 **Jin W,** Gao MQ, Lin ZW, Yang DX. Multiple biomarkers of colorectal tumor in a differential diagnosis model: a quantitative study. *World J Gastroenterol* 2004; **10**: 439-442

29 **Wilcox S,** Parra-Medina D, Thompson-Robinson M, Will J. Nutrition and physical activity interventions to reduce cardiovascular disease risk in health care settings: a quantitative review with a focus on women. *Nutr Rev* 2001; **59**: 197-214

30 **Chen K.** A quantitative method for mass screening of colorectal cancer. *Zhejiang Yike Daxue Xuebao* 1988; **17**: 49-52

31 **Chen K,** Zhu YM, Fang SY, Chen YY. Quantitative method for mass screening of coronary heart disease (CHD) and stroke. *Jibing Kongzhi Zazhi* 1997; **1**: 15-17

32 **Zhou BS,** He AG, Liu KL, Sun J, Shi HL, Gu DY, Gu HC. Quantitative scored lung cancer risk assessment. *Zhongguo Weisheng Tongji* 1996; **13**: 19-22

Edited by Wang XL and Kumar M