Factors associated with gastric adenocarcinoma and dysplasia in patients with chronic gastritis: a population-based study

Jie Xing[1], Li Min[1], Shengtao Zhu[1], Hao Zhang[2], Yu Zhao[1], Hengcun Li[1], Zheng Zhang[1], Peng Li[1], Shutian Zhang[1]

[1]Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Disease, Beijing Digestive Disease Center, Beijing Key Laboratory for Precancerous Lesion of Digestive Disease, Beijing 100050, China; [2]Department of Gastroenterology, Jizhong Energy Fengfeng Group Hospital, Handan 056200, China

Correspondence to: Prof. Shutian Zhang, MD, PhD. Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Disease, Beijing Digestive Disease Center, Beijing Key Laboratory for Precancerous Lesion of Digestive Disease, Beijing 100050, China. Email: zhangshutian@ccmu.edu.cn.

Abstract

Objective: Gastric cancer (GC) is one of the leading causes of death in China and other Asian countries. Recently, gastric endoscopy has become the main approach for GC screening, but the identification of high-risk individuals remains a challenge in GC screening programs.

Methods: There were 7,302 patients with chronic gastritis involved in this study. Endoscopic examinations were performed, and their demographic characteristics and lifestyle data were collected. Each possible associated factor of GC/premalignant and precursor lesions was evaluated by univariate and multivariate logistic regressions. Nomograms were used for visualization of those models, and receiver operating characteristic (ROC) curve analysis was used to present the predictive accuracy.

Results: We detected 8 (0.11%) gastric adenocarcinomas, 17 (0.23%) dysplasia cases, 14 (0.19%) hyperplasia cases, 52 (0.71%) intestinal metaplasia cases, 217 (2.97%) inflammatory lesions, 141 (1.93%) gastric ulcers, 10 (0.14%) atrophic gastritis cases, 1,365 (18.69%) erosive gastritis cases, and 5,957 (81.58%) superficial gastritis cases in 7,302 patients. The age (P<0.001), gender (P=0.086), labor intensity (P=0.018) and leek food intake (P=0.143) were identified as independent predictive factors of GC/premalignant lesions possibility. The corresponding nomogram exhibited an area under the curve (AUC) [95% confidence interval (95% CI)] of 0.82 (0.74–0.89) for the modeling group and 0.80 (0.75–0.85) for the validation group. The age (P=0.002), gender (P=0.024), smoking (P=0.002) and leek food intake (P=0.039) were independent predictive factors of precursor lesions possibility. The corresponding nomogram exhibited an AUC (95% CI) of 0.62 (0.60–0.65) for the modeling group and 0.61 (0.59–0.63) for the validation group.

Conclusions: We identified several potential associated factors and provided a preclinical nomogram with the potential to predict the possibility of GC/premalignant and precursor lesions.

Keywords: Risk; GC; dysplasia; gastritis; population-based study

Introduction

Gastric cancer (GC) is one of the most commonly diagnosed carcinomas worldwide with 680,000 new cases and 754,000 deaths in 2012 (1). GC has an especially high incidence and mortality in East Asia and China, accounting for approximately 40% of GC cases and accompanied by a poor nutrition status and unhealthy living habits of the

342

Xing et al. Predictors of GC and dysplasia

residents in high-risk areas (2). Although gastric endoscopy has been used in the screening for GC and has a decreasing trend in mortality, GC remains the main cancer burden in some areas (3).

GC is highly aggressive and usually involves lymph nodes and distant organ metastasis (4). Most patients are diagnosed at a late stage and have a very poor prognosis (5). Many studies have shown that the overall 5-year survival of GC has a strong relationship with the stage at detection and treatment (6). Kim *et al.* reported that the 5-year survival rates of GC decreased dramatically when it was not diagnosed at a very early stage; therefore, proper screening management for GC is important (7). Increasing the early detection and treatment rate of GC would be very beneficial for GC patients (8,9).

Recently, for GC, premalignant and precursor lesions have been studied in detail. Risk factors, such as *Helicobacter pylori* (HP) infection, chronic atrophic gastritis (CAG) and intestinal metaplasia (IM), are widely recognized and adopted in GC screening guidelines (10). However, some studies have also showed that HP eradication did not completely reduce the incidence of GC (11); on the other hand, many other risk factors have been proposed, such as the age, gender, smoking, drinking, other dietary factors, etc., but there is no consensus (12). Due to the lack of clinical and epidemiological evidence to select eligible people with high risk, GC screening has nearly completely depended on endoscopy (13). Recent GC screening programs have only selected patients by HP and CAG, especially in China (14), resulting in low compliance with screening and waste of medical resources.

In this study, we established a cohort of 7,302 chronic gastritis patients in a single medical center who were screened for GC as well as its premalignant and precursor lesions by endoscopy; we then collected epidemiological data. Here, we identified several associated factors of GC/premalignant and precursor lesions and established a method to select eligible populations for endoscopic screening based on demographic data, lifestyle factors, eating habits, and psychological factors.

Materials and methods

Participant inclusion/exclusion criteria

A total of 7,497 chronic gastritis patients in Jizhong Energy Fengfeng Group Hospital from Handan City, Hebei Province, were enrolled in this study during October 2010 to September 2011, and all were 40–70 years old. All 7,497

patients provided informed consent and 7,484 completed the epidemiology questionnaire. General physical examination was conducted, and the reports showed that 182 people were excluded because of: 1) iodine allergy; 2) diagnosed with other tumors; 3) pregnancy; 4) hyperthyroidism; 5) lactation; or 6) psychosis. Based on the questionnaire and health examination, 7,302 participants were recruited, and they all successfully underwent a gastrointestinal endoscopy examination (*Figure 1*). All procedures in this study had been approved by the Ethics Committee of Beijing Friendship Hospital.

Performance of endoscopy and its quality control

All endoscopic procedures (examinations and therapies) were performed by experienced endoscopists. Before the endoscopic procedure, all participants were required to have an empty stomach for 8 h and stop anticoagulant drugs for 1 week. We used iodine staining to facilitate the discovery of lesions, and any lesions were biopsied, including carcinoma, hyperplasia, dysplasia, ulcers, inflammatory lesions, and others. Also, we examined 5% of normal gastric mucosa as negative controls. The diagnostic results were determined by two pathologists based on the diagnostic criteria of the World Health Organization.

Acquisition and processing of data

We acquired data through questionnaires of demographics, eating habits, lifestyle factors, psychological factors, and others. Smoking status was defined by four subgroups: "yes" means smoking for more than 6 months and ≥1 cigarettes per day; "occasionally" means <1 cigarette/d or less than 6 months; "refrained" means stopped smoking for
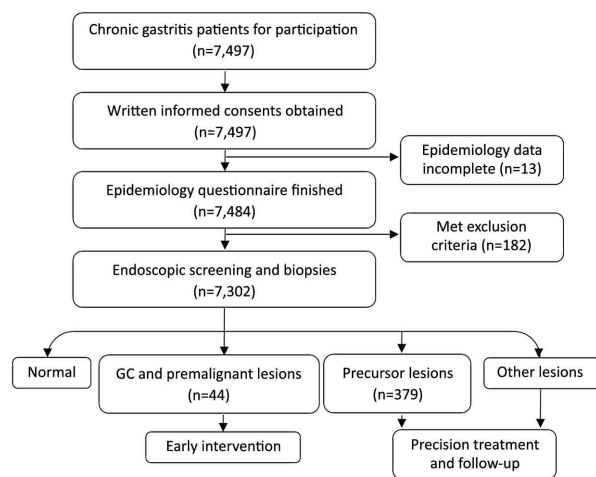

Figure 1 Flow chart of patient inclusion and exclusion.

more than 6 months; and "never" means never smoke. Eating habits (including leek food intake, fresh vegetable consumption, etc.) were defined by two subgroups: "yes" means ≥3 per month; and "rarely or never" means never or <3 per month. Data on different stomach lesions were collected by endoscopic examinations, and the consensus of two pathologists provided the pathological diagnosis data. We imported all questionnaires and relative information by two independent researchers into the database.

Statistical analysis and modeling

R Software (Version 3.3.1; Rocket Software, Inc. Waltham, MA, USA) was used to analyze all statistics with logistic regression models. Packages "dplyr", "plyr", and "reshape2" were used for data cleaning. Packages "rms" and "ordinal" were used for nomogram building. Package "ggplot2" was used for visualization. We conducted univariate analysis to identify the risk factors of gastric lesions, and displayed the odds ratio (OR) with 95% confidence interval (95% CI). We use multivariate regression analysis to identify the independent influence factors. Nomograms were established based on the results by half of the cases which sampled by the Monte Carlo method. The other half of the cases were used as a validation set. The receiver operating characteristic (ROC) curve was used to graphically display the predictive accuracy of nomograms. The area under the curve (AUC) for validation was used to assess the nomogram accuracy.

Results

Baseline data of chronic gastritis patient population

Overall, 7,302 chronic gastritis patients successfully underwent gastric endoscopy and completed epidemiology questionnaires. Detailed data of the demographics and lifestyle factors were obtained. We detected 8 (0.11%) gastric carcinomas, and all were adenocarcinomas. Seventeen (0.23%) dysplasia and 14 (0.19%) hyperplasia cases were also detected. There were 52 (0.71%) cases with intestinal metaplasia, 217 with inflammatory lesions (2.97%), 141 (1.93%) with gastric ulcers, 10 (0.14%) with atrophic gastritis, 1,365 (18.69%) with erosive gastritis, and 5,957 (81.58%) with superficial gastritis (*Table 1*).

Among all 7,302 patients, 6,574 (90.03%) were between 41 and 60 years old, 5,008 (68.58%) were male, 5,246 (71.84%) were workers, and 5,229 (71.61%) earned ¥2,001–3,000 ($286–428) per month. The demographic baseline data of this population are also displayed in *Table 2*.

Table 1 Gastric lesions in patients (N=7,302)

Lesions	n	%
Carcinoma	8	0.11
Dysplasia	17	0.23
Hyperplasia	14	0.19
Intestinal metaplasia	52	0.71
Inflammatory lesions	217	2.97
Gastric ulcer	141	1.93
Atrophic gastritis	10	0.14
Erosive gastritis	1,365	18.69
Superficial gastritis	5,957	81.58

Table 2 Demographic baseline data of chronic gastritis patient population

Characteristics	Case No.	%
Age (year)		
≤40	473	6.48
41–50	4,293	58.79
51–60	2,281	31.24
>60	255	3.49
Gender		
Male	5,008	68.58
Female	2,294	31.42
Education		
College and above	824	11.28
High school	2,630	36.02
Middle school	3,356	45.96
Primary school	492	6.74
Occupation		
Worker	5,246	71.84
Peasant	102	1.40
Officer	968	13.26
Service	255	3.49
Retired	629	8.61
Others	102	1.40
Labor intensity		
Light	3,747	51.31
Moderate	2,553	34.96
Heavy	1,002	13.72
Income (¥)		
≤2,000	1,597	21.87
2,001–3,000	5,229	71.61
3,001–4,000	1,751	23.98
>4,000	1,425	19.52

344

Xing et al. Predictors of GC and dysplasia

Detection of associated factors for GC/premalignant lesions

To identity the associated factors for GC and premalignant lesions, we performed univariate logistic regression analyses of GC/premalignant lesions (*Table 3* and *Supplementary Table S1*). We found that age (P<0.001), occupation (P=0.030), labor intensity (P=0.006), tea consumption (P=0.014), and water source (P=0.050) were all significantly associated with GC and premalignant lesion incidence, while gender (P=0.077), tumor history (P=0.093), high temperature food intake (P=0.063), and leek food intake (P=0.079) were at a marginal level of association.

However, in multivariate analysis, only the age (P<0.001), gender (P=0.086), labor intensity (P=0.018) and leek food intake (P=0.143) were independent predictive factors of GC/premalignant lesions possibility (*Table 3*).

Table 3 Univariate and multivariate analyses of GC and premalignant lesions associated factors

Characteristics	Univariate*			Multivariate		
	OR	95% CI	P	OR	95% CI	P
Age (year)			<0.001			<0.001
41–50 *vs.* ≤40	$1.4E^6$	NA		$1.2E^6$	NA	
51–60 *vs.* ≤40	$5.3E^6$	NA		$3.9E^6$	NA	
>60 *vs.* ≤40	$1.4E^7$	NA		$9.7E^6$	NA	
Gender			0.077			0.086
Female *vs.* male	0.42	(0.12, 1.09)		0.38	(0.11, 1.04)	
Occupation			0.030			
Peasant *vs.* worker	8.72	(1.35, 32.5)				
Officer *vs.* worker	1.36	(0.31, 42.8)				
Service *vs.* worker	3.45	(0.53, 12.7)				
Retired *vs.* worker	4.20	(1.46, 1.09)				
Others *vs.* worker	$3.8E^{-6}$	NA				
Labor intensity			0.006			0.018
Moderate *vs.* light	0.15	(0.02, 0.53)		0.19	(0.03, 0.69)	
Heavy *vs.* light	0.79	(0.23, 2.10)		1.02	(0.29, 2.89)	
Income (¥)			0.390			
2,001–3,000 *vs.* ≤2,000	0.42	(0.14, 1.17)				
3,001–4,000 *vs.* ≤2,000	0.51	(0.15, 1.47)				
>4,000 *vs.* ≤2,000	0.62	(0.19, 1.80)				
Smoking			0.147			
Occasionally *vs.* Yes	0.66	(0.10, 2.33)				
Refrained *vs.* Yes	0.36	(0.02, 1.79)				
Never *vs.* Yes	0.38	(0.14, 0.89)				
Tumor history			0.093			
No *vs.* Yes	2.32	(0.88, 7.95)				
Other disease history			0.302			
No *vs.* Yes	2.00	(0.59, 12.45)				
Tea consumption			0.014			
No *vs.* Yes	0.33	(0.15, 0.79)				
Eating postures			0.629			
Sit *vs.* squat	$1.1E^6$	NA				
Stand *vs.* squat	1.00	NA				

Table 3 (*continued*)

Characteristics	Univariate*			Multivariate		
	OR	95% CI	P	OR	95% CI	P
High temperature food intake			0.063			
No *vs*. Yes	2.27	(0.96, 6.24)				
Leek food consumption			0.079			0.143
Rarely or never *vs*. Yes	0.24	(0.01, 1.14)		0.291	(0.02, 1.39)	
Fresh vegetable consumption			0.108			
Rarely or never *vs*. Yes	9.15	(0.50, 45.32)				
Water source			0.050			
Natural *vs*. others	3.44	(1.00, 9.08)				
Foreign residency			0.951			
No *vs*. Yes	0.94	(0.20, 16.82)				

GC, gastric cancer; OR, odds ratio; 95% CI, 95% confidence intervals; NA, not available; *, partial results of univariate of GC and premalignant lesions associated factors, full table of this part was displayed in *Supplementary Table S1*.

Detection of associated factors of precursor lesions

To identity the associated factors of GC precursor lesions, we performed univariate logistic regression analyses of precursor lesions (*Table 4* and *Supplementary Table S1*). We found that age (P=0.002), gender (P=0.016), income (P=0.018), and smoking (P<0.001) were significantly associated with the GC and premalignant lesion incidence, while leek food intake was at a marginal level (P=0.090).

However, in multivariate analysis, only age (P=0.002), gender (P=0.024), smoking (P=0.002) and leek food intake (P=0.039) were independent predictive factors of precursor lesions possibility (*Table 4*).

Nomograms for predicting gastric lesions

All independent GC/premalignant and precursor lesion associated factors (age, gender, etc.) were included to develop predictive nomograms for predicting those diseases, which would provide very helpful information in the clinical evaluation and selection of eligible people for GC screening.

Table 4 Univariate and multivariate analyses of precursor lesions associated factors

Characteristics	Univariate*			Multivariate		
	OR	95% CI	P	OR	95% CI	P
Age (year)			0.002			0.002
41–50 *vs*. ≤40	2.10	(1.13, 4.45)		2.03	(1.09, 4.32)	
51–60 *vs*. ≤40	2.73	(1.46, 5.85)		2.63	(1.39, 5.62)	
>60 *vs*. ≤40	3.45	(1.53, 8.26)		3.73	(1.65, 8.96)	
Gender			0.016			0.024
Female *vs*. male	0.73	(0.56, 0.95)		1.06	(0.72, 1.56)	
Occupation			0.640			
Peasant *vs*. worker	1.19	(0.42, 2.68)				
Officer *vs*. worker	1.03	(0.72, 1.43)				
Service *vs*. worker	0.95	(0.46, 1.72)				
Retired *vs*. worker	1.28	(0.87, 1.84)				
Others *vs*. worker	0.46	(0.08, 1.47)				
Labor intensity			0.750			
Moderate *vs*. light	1.03	(0.80, 1.31)				
Heavy *vs*. light	0.89	(0.61, 1.26)				

Table 4 (*continued*)

Characteristics	Univariate*			Multivariate		
	OR	95% CI	P	OR	95% CI	P
Income (¥)			0.018			
2,001–3,000 *vs.* ≤2,000	0.70	(0.53, 0.94)				
3,001–4,000 *vs.* ≤2,000	0.64	(0.46, 0.88)				
>4,000 *vs.* ≤2,000	0.63	(0.44, 0.90)				
Smoking			<0.001			0.002
Occasionally *vs.* Yes	0.73	(0.45, 1.11)		0.73	(0.45, 1.12)	
Refrained *vs.* Yes	0.59	(0.34, 0.95)		0.54	(0.31, 0.87)	
Never *vs.* Yes	0.58	(0.46, 0.75)		0.55	(0.27, 0.78)	
Tumor history			0.245			
No *vs.* Yes	1.16	(0.90, 1.51)				
Other disease history			0.470			
No *vs.* Yes	1.13	(0.82, 1.60)				
Tea consumption			0.512			
No *vs.* Yes	0.90	(0.67, 1.23)				
Eating postures			0.460			
Sit *vs.* squat	0.98	(0.44, 2.79)				
Stand *vs.* squat	$3.8E^{-6}$	NA				
High temperature food intake			0.263			
No *vs.* Yes	1.14	(0.91, 1.45)				
Leek food consumption			0.090			0.039
Rarely or never *vs.* Yes	1.30	(0.96, 1.73)		1.38	(1.02, 1.85)	
Fresh vegetable consumption			0.240			
Rarely or never *vs.* Yes	2.21	(0.53, 6.23)				
Water source			0.660			
Natural *vs.* others	1.12	(0.66, 1.77)				
Foreign residency			0.121			
No *vs.* Yes	0.66	(0.41, 1.13)				

GC, gastric cancer; OR, odds ratio; 95% CI, 95% confidence intervals; *, partial results of univariate of precursor lesions associated factors, full table of this part was displayed in *Supplementary Table S1*.

Half of the cases were sampled by the Monte Carlo method to generate predictive nomograms (*Figure 2A*, *D*). For each case, we assigned a point for each factor, and then summed them to generate a total score. A corresponding predicted probability of certain lesions was calculated from the nomogram. ROC curves were used to evaluate the predictive accuracy of the nomograms. For prediction of GC/premalignant lesions, the AUC (95% CI) was 0.82 (0.74–0.89) for the modeling group (*Figure 2B*) and 0.80 (0.75–0.85) for the validation group (*Figure 2C*). In the prediction of precursor lesions, the AUC (95% CI) was 0.62 (0.60–0.65) for the modeling group (*Figure 2E*) and 0.61 (0.59–0.63) for the validation group (*Figure 2F*).

Discussion

Many independent studies have suggested that identification of GC at an early stage could reduce GC mortality and substantially decrease medical burden (15-17). Additionally, detection of precancerous and precursor lesions could also reduce both the incidence and mortality of GC (18,19). Gastric endoscopy is the main approach to screen for GC and its precancerous lesions. Recently, most GC endoscopy screening programs were conducted in a target population that was roughly defined by specific associated factors, such as age and family history (20-22).

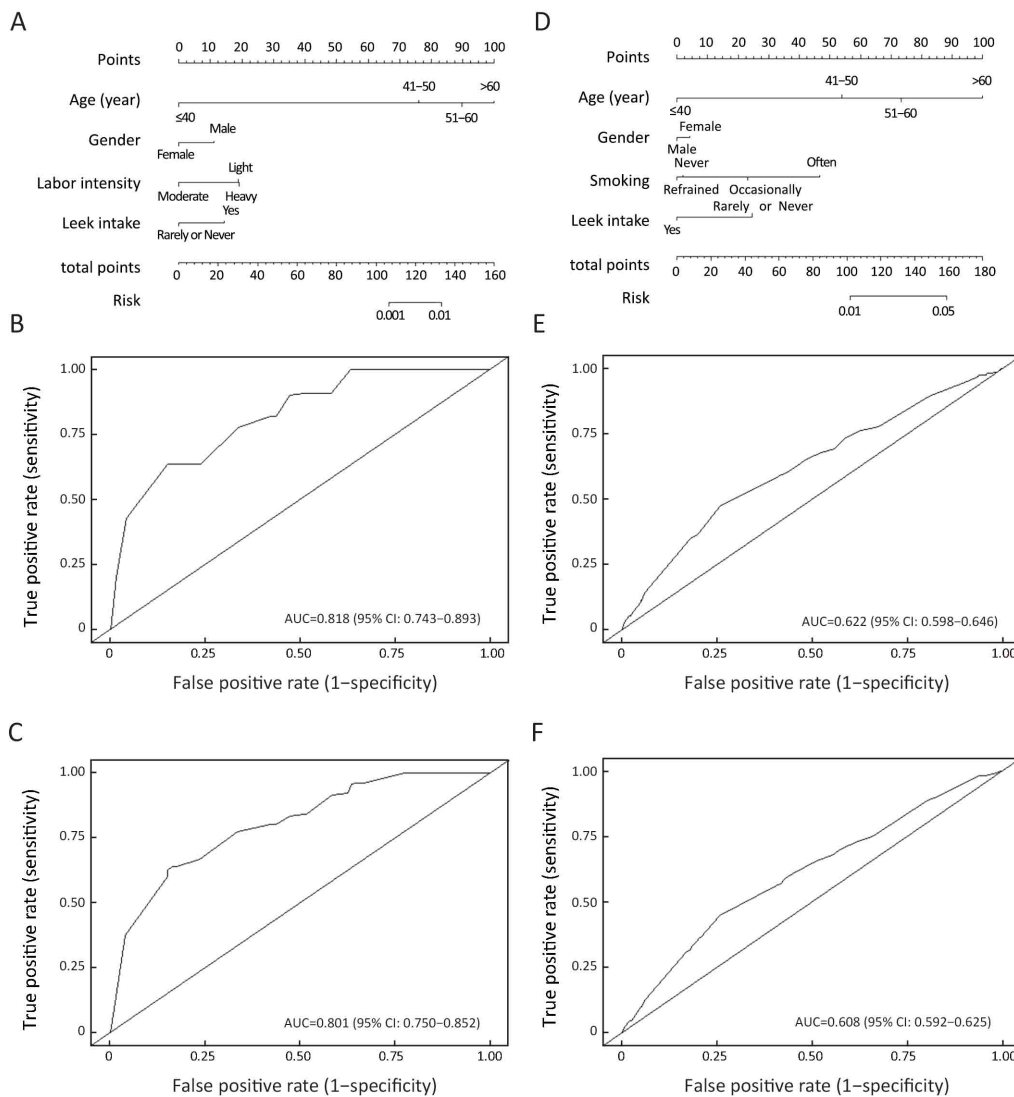Inefficient pre-selection of high-risk people

Figure 2 Nomograms for predicting gastric cancer (GC)/premalignant and precursor lesions possibility. (A) Predictive nomogram of GC/premalignant lesions (adenocarcinoma, dysplasia and intestinal metaplasia); (B) Receiver operating characteristic (ROC) curve of a nomogram model to predict the GC/premalignant lesions possibility (modeling data); (C) ROC curve of a nomogram model to predict the GC/premalignant lesions possibility (validation data); (D) Predictive nomogram of precursor lesions (erosive gastritis, atrophic gastritis and gastric ulcer); (E) ROC curve of the nomogram model to predict the precursor lesion possibility (modeling data); (F) ROC curve of the nomogram model to predict the precursor lesion possibility (validation data).

demonstrated low cost-efficiency in most GC screening programs. The low true positive rate in screening was also influenced by very weak compliance according to our experience. Systematic evaluation of the possible associated factors of GC/premalignant and precursor lesions would further our understanding of GC and facilitate endoscopists' decision making in terms of whether to perform endoscopy examination.

In this study, it was not surprising that GC/premalignant

and precursor lesions largely shared independent associated factors, considering the natural historical origination of GC (23). Age was greatly associated with both GC/premalignant and precursor lesion incidence, which may be caused by DNA injury accumulation in the elderly (24). Decreased immunity in the elderly could also explain the general increase in most diseases (25). Additionally, we found that men were more likely to develop GC/premalignant lesions, while women were more likely to be

affected by precursor lesions. This phenomenon could be explained by potential confounding factors, such as different lifestyles between men and women (men are more addicted to smoking and drinking, while women are more likely to go on a drastic diet) (26). Previous studies suggested that leek consumption could reduce the gastrointestinal cancer risk in the general population. However, we found that leek food intake was positively associated with GC/premalignant incidence, while it was negatively associated with the precursor lesion incidence. The exact role of leek food intake in GC cancer development requires more, higher level evidence.

Nomograms exhibited very promising potential in clinical trial design and interpretation; also, they were widely adopted in prognostic models (27-29). However, they were rarely used in primary health screening (30,31), and they were never used in GC screening programs. In this study, nomograms were used as a visualization tool to display the prediction model based on logistic regression. Here, we established a nomogram-based method to evaluate the possibility of GC/premalignant and precursor lesions in chronic gastritis patients, which is the first nomogram to predict the individual GC risk based on demographic and epidemiological data. The AUC of our model ranged from 0.61 to 0.82, suggesting that a nomogram would be a very promising tool in individual GC risk assessment, especially in primary health screening programs.

There were also several shortcomings in this study. First, limited by the condition in the local medical center, we did not examine the HP status, which may be a very important GC associated factor. Second, because only demographic and lifestyle information were collected to build disease associated models, no biochemistry or hematology tests were involved in this study. Third, our nomograms were based on a retrospective single centered cohort. Nevertheless, we did not aim to provide a perfect prediction of GC/premalignant and precursor lesions; instead, our goal was to deliver an easy to use tool for rough filtering of an eligible population for GC screening that is better than the currently available strategies. Additionally, there were only 8 carcinomas detected in our population, which is much lower than expected. The composition of participants and their age distribution might partially explain this outcome. Most of the patients from Jizhong Energy Fengfeng Group Hospital were workers in Jizhong Energy Fengfeng Group Co. Ltd., aged under 60 years (before retirement). Considering the elder people (>60 years old) have much higher GC risk than others, it is not surprising we obtained a low carcinoma detection rate. Thus, we could only investigate GC along with its premalignant lesions.

Conclusions

We provided a systematic evaluation of possible associated factors for GC/premalignant and precursor lesions based on a large population of chronic gastritis patients and then generated nomograms to predict the disease possibility of every individual. A multi-centered prospective study with a larger population would be expected to provide a more reliable estimation in the future.

Acknowledgements

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J Clin 2015;65:87-108.

2. Chen W, Zheng R, Zuo T, et al. National cancer incidence and mortality in China, 2012. Chin J Cancer Res 2016;28:1-11.

3. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127:2893-917.

4. Jin X, Zhu Z, Shi Y. Metastasis mechanism and gene/protein expression in gastric cancer with distant organs metastasis. Bull Cancer 2014;101:E1-12.

5. Van Cutsem E, Sagaert X, Topal B, et al. Gastric cancer. Lancet 2016;388:2654-64.

6. Khanderia E, Markar SR, Acharya A, et al. The influence of gastric cancer screening on the stage at diagnosis and survival: a Meta-analysis of comparative

studies in the Far East. J Clin Gastroenterol 2016;50:190-7.

7. Kim YG, Kong SH, Oh SY, et al. Effects of screening on gastric cancer management: comparative analysis of the results in 2006 and in 2011. J Gastric Cancer 2014;14:129-34.

8. Jung DH, Hwang S, Song GW, et al. Survival benefit of early cancer detection through regular endoscopic screening for de novo gastric and colorectal cancers in Korean liver transplant recipients. Transplant Proc 2016;48:145-51.

9. Wei WQ, Yang CX, Lu SH, et al. Cost-benefit analysis of screening for esophageal and gastric cardiac cancer. Chin J Cancer 2011;30:213-8.

10. Karimi P, Islami F, Anandasabapathy S, et al. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev 2014;23:700-13.

11. Asaka M, Kato M, Sakamoto N. Roadmap to eliminate gastric cancer with Helicobacter pylori eradication and consecutive surveillance in Japan. J Gastroenterol 2014;49:1-8.

12. Choi IJ. Endoscopic gastric cancer screening and surveillance in high-risk groups. Clin Endosc 2014;47:497-503.

13. Choi KS, Jun JK, Suh M, et al. Effect of endoscopy screening on stage at gastric cancer diagnosis: results of the National Cancer Screening Programme in Korea. Br J Cancer 2015;112:608-12.

14. Chen XZ, Schöttker B, Castro FA, et al. Association of helicobacter pylori infection and chronic atrophic gastritis with risk of colonic, pancreatic and gastric cancer: A ten-year follow-up of the ESTHER cohort study. Oncotarget 2016;7:17182-93.

15. Suh YS, Yang HK. Screening and early detection of gastric cancer: East versus West. Surg Clin North Am 2015;95:1053-66.

16. Jang HJ, Choi MH, Shin WG, et al. Is annual endoscopic surveillance necessary for the early detection of gastric remnant cancer in Korea? A retrospective multi-center study. Hepatogastroenterology 2014;61:1283-6.

17. Paterson HM, McCole D, Auld CD. Impact of open-access endoscopy on detection of early oesophageal and gastric cancer 1994-2003: population-based study. Endoscopy 2006;38:503-7.

18. Amal H, Leja M, Funka K, et al. Detection of precancerous gastric lesions and gastric cancer through exhaled breath. Gut 2016;65:400-7.

19. Kang JH, Lim YJ, Kang JH, et al. Prevalence of precancerous conditions and gastric cancer based upon the national cancer screening program in Korea for 7 years, single center experience. Gastroenterol Res Pract 2015;2015:571965.

20. Kim GH, Liang PS, Bang SJ, et al. Screening and surveillance for gastric cancer in the United States: Is it needed? Gastrointest Endosc 2016;84:18-28.

21. Yeh JM, Hur C, Ward Z, et al. Gastric adeno-carcinoma screening and prevention in the era of new biomarker and endoscopic technologies: a cost-effectiveness analysis. Gut 2016;65:563-74.

22. Park CH, Kim EH, Jung DH, et al. The new modified ABCD method for gastric neoplasm screening. Gastric Cancer 2016;19:128-35.

23. Hohenberger P, Gretschel S. Gastric cancer. Lancet 2003;362:305-15.

24. Tomasetti C, Li L, Vogelstein B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. Science 2017;355:1330-4.

25. Yaqoob P. Ageing alters the impact of nutrition on immune function. Proc Nutr Soc 2016;1-5.

26. Trovato FM, Martines GF, Brischetto D, et al. Fatty liver disease and lifestyle in youngsters: diet, food intake frequency, exercise, sleep shortage and fashion. Liver Int 2016;36:427-33.

27. Kim SK, Chai YJ, Park I, et al. Nomogram for predicting central node metastasis in papillary thyroid carcinoma. J Surg Oncol 2017;115:266-72.

28. Zhang J, Li X, Huang R, et al. A nomogram to predict the probability of axillary lymph node metastasis in female patients with breast cancer in China: A nationwide, multicenter, 10-year epidemiological study. Oncotarget 2017;8:35311-25.

29. Zhou ML, Wang L, Wang JZ, et al. Validation of the Memorial Sloan Kettering Cancer Center nomogram to predict disease-specific survival in a Chinese gastric cancer population receiving postoperative chemoradiotherapy after an R0 resection. Oncotarget 2016;7:64757-65.

30. Bass JL, Bhatia A, Boas FE, et al. Validation of a body mass index nomogram for children as an obesity

screening tool in young children. Clin Pediatr (Phila) 2006;45:718-24.

31. Rozendaal L, Groenink M, Naeff MS, et al. Marfan syndrome in children and adolescents: an adjusted nomogram for screening aortic root dilatation. Heart 1998;79:69-72.

Table S1 Univariate analysis of gastric lesions associated factors

Characteristics	Precursor lesions			Premalignant/malignant lesions		
	OR	95% CI	P	OR	95% CI	P
Age (year)			<0.002			<0.001
41–50 vs. ≤40	2.10	(1.13, 4.45)		1.40^6	($6.17E^{-15}$, NA)	
51–60 vs. ≤40	2.73	(1.46, 5.85)		5.28^6	($2.33E^{-14}$, NA)	
>60 vs. ≤40	3.45	(1.53, 8.26)		1.36^7	($6.02E^{-14}$, NA)	
Gender			<0.016			0.077
Female vs. male	0.73	(0.56, 0.95)		0.42	(0.12, 1.09)	
Race			0.218			0.156
Others vs. Han	2.04	(0.61, 5.06)		6.55	(0.36, 32.06)	
Married			0.624			0.475
Others vs. married	1.30	(0.39, 3.17)		2.49^{-6}	(NA, 2.61)	
Education			0.393			0.536
High school vs. college	1.10	(0.74, 1.69)		0.63	(0.16, 2.97)	
Middle school vs. college	1.20	(0.82, 1.83)		1.06	(0.34, 4.65)	
Primary school vs. college	1.54	(0.90, 2.62)		1.68	(0.31, 9.01)	
Occupation			0.640			0.030
Peasant vs. worker	1.19	(0.42, 2.68)		8.72	(1.35, 32.5)	
Officer vs. worker	1.03	(0.72, 1.43)		1.36	(0.31, 4.28)	
Service vs. worker	0.95	(0.46, 1.72)		3.45	(0.53, 12.70)	
Retired vs. worker	1.28	(0.87, 1.84)		4.20	(1.46, 1.09)	
Others vs. worker	0.46	(0.08, 1.47)		3.77^{-6}	(NA, $5.32E^{10}$)	
Labor intensity			0.750			0.006
Moderate vs. light	1.03	(0.80, 1.31)		0.15	(0.02, 0.53)	
Heavy vs. light	0.89	(0.61, 1.26)		0.79	(0.23, 2.10)	
Income (¥)			0.018			0.390
2,001–3,000 vs. ≤2,000	0.70	(0.53, 0.94)		0.42	(0.14, 1.17)	
3,001–4,000 vs. ≤2,000	0.64	(0.46, 0.88)		0.51	(0.15, 1.47)	
>4,000 vs. ≤2,000	0.63	(0.44, 0.90)		0.62	(0.19, 1.80)	
Mining exposure			<0.527			0.445
Yes vs. No	1.08	(0.86, 1.35)		0.73	(0.32, 1.62)	
Mining exposure time (hours/week)			<0.482			0.573
1–40 vs. 0	1.11	(0.74, 1.61)		0.71	(0.11, 2.52)	
41–60 vs. 0	1.18	(0.90, 1.53)		0.93	(0.35, 2.21)	
>60 vs. 0	0.88	(0.58, 1.30)		0.31	(0.02, 1.51)	
Dust exposure			0.982			0.192
Yes vs. No	1.00	(0.79, 1.25)		0.59	(0.26, 1.30)	
Gas exposure			0.287			0.124
Yes vs. No	0.88	(0.70, 1.11)		0.52	(0.20, 1.19)	
Smoking			<0.001			0.147
Occasionally vs. Yes	0.73	(0.45, 1.11)		0.66	(0.10, 2.33)	
Refrained vs. Yes	0.59	(0.34, 0.95)		0.36	(0.02, 1.79)	
Never vs. Yes	0.58	(0.46, 0.75)		0.38	(0.14, 0.89)	

Table S1 (*continued*)

Characteristics	Precursor lesions			Premalignant/malignant lesions		
	OR	95% CI	P	OR	95% CI	P
Drinking			0.243			0.526
Occasionally *vs*. Yes	0.85	(0.64, 1.13)		9.67^{-1}	(0.38, 2.40)	
Refrained *vs*. Yes	0.84	(0.63, 1.12)		6.71^{-1}	(0.23, 1.81)	
Never *vs*. Yes	1.42	(0.77, 2.45)		7.95^{-7}	(NA, 1.45E^{12})	
Smoking exposure			0.139			0.441
Yes *vs*. No	0.81	(0.61, 1.07)		0.69	(0.29, 1.89)	
Smoking time >15 min			0.641			0.588
Yes *vs*. No	1.06	(0.83, 1.35)		1.26	(0.53, 2.79)	
Disease history			0.680			0.615
No *vs*. Yes	1.05	(0.83, 1.32)		1.23	(0.56, 2.82)	
Heart disease			0.604			0.830
No *vs*. Yes	0.90	(0.63, 1.35)		1.17	(0.34, 7.29)	
Diabetes			0.606			0.211
No *vs*. Yes	1.20	(0.63, 2.67)		1.11^{6}	(1.17E^{-12}, NA)	
Hypertension			0.934			0.394
No *vs*. Yes	0.99	(0.76, 1.30)		1.56	(0.59, 5.34)	
Digestive disease			0.442			0.718
No *vs*. Yes	1.14	(0.82, 1.62)		1.24	(0.43, 5.25)	
Respiratory disease			0.639			0.263
No *vs*. Yes	0.91	(0.63, 1.37)		0.52	(0.20, 1.78)	
Tumor history			0.245			0.093
No *vs*. Yes	1.16	(0.90, 1.51)		2.32	(0.88, 7.95)	
Other disease history			0.470			0.302
No *vs*. Yes	1.13	(0.82, 1.60)		2.00	(0.59, 12.45)	
Tea consumption			0.512			0.014
No *vs*. Yes	0.90	(0.67, 1.23)		0.33	(0.15, 0.79)	
Eating postures			0.460			0.629
Sit *vs*. squat	0.98	(0.44, 2.79)		1.10^{6}	(9.95^{-20}, NA)	
Stand *vs*. squat	3.86E^{-6}	(8.04E^{-70}, 3.17E^{-103})		1.00	(6.02^{-31}, 1.66^{30})	
High temperature food intake			0.263			0.063
No *vs*. Yes	1.14	(0.91, 1.45)		2.27	(0.96, 6.24)	
Fast eating			0.206			0.111
No *vs*. Yes	0.86	(0.68, 1.08)		0.52	(0.21, 1.16)	
Salted food consumption			0.581			0.140
Rarely or never *vs*. Yes	0.94	(0.74, 1.19)		0.55	(0.25, 1.22)	
Fried food consumption			0.367			0.234
Rarely or never *vs*. Yes	1.13	(0.87, 1.49)		1.84	(0.70, 6.30)	
Chilli food consumption			0.526			0.733
Rarely or never *vs*. Yes	0.92	(0.72, 1.19)		1.17	(0.49, 3.22)	
Moldy food consumption			0.766			0.527
No *vs*. Yes	1.23	(0.38, 7.54)		4.01^{5}	(1.31^{-16}, NA)	

Table S1 (*continued*)

Characteristics	Precursor lesions			Premalignant/malignant lesions		
	OR	95% CI	P	OR	95% CI	P
Leek food consumption			0.090			0.079
Rarely or never *vs.* Yes	1.30	(0.96, 1.73)		0.24	(0.01, 1.14)	
Bean food consumption			0.912			0.570
Rarely or never *vs.* Yes	1.02	(0.74, 1.37)		0.71	(0.17, 2.07)	
Meat/egg/milk consumption			0.699			0.573
Rarely or never *vs.* Yes	1.09	(0.68, 1.66)		0.59	(0.03, 2.79)	
Dry food consumption			0.815			0.355
Rarely or never *vs.* Yes	1.03	(0.81, 1.31)		0.69	(0.31, 1.55)	
Maize consumption			0.814			0.811
Rarely or never *vs.* Yes	1.03	(0.81, 1.30)		1.10	(0.48, 2.43)	
Fresh vegetable consumption			0.240			0.108
Rarely or never *vs.* Yes	2.21	(0.53, 6.23)		9.15	(0.50, 45.32)	
Fresh fruit consumption			0.582			0.585
Rarely or never *vs.* Yes	1.07	(0.85, 1.34)		0.80	(0.36, 1.77)	
Sauerkraut consumption			0.623			0.405
No *vs.* Yes	1.09	(0.78, 1.55)		0.65	(0.26, 1.95)	
Water source			0.660			0.050
Natural *vs.* others	1.12	(0.66, 1.77)		3.44	(1.00, 9.08)	
Foreign residency			0.121			0.951
No *vs.* Yes	0.66	(0.41, 1.13)		0.94	(0.20, 16.82)	

OR, odds ratio; 95% CI, 95% confidence intervals; NA, not available.