

Development and validation of a questionnaire-based risk scoring system to identify individuals at high risk for gastric cancer in Chinese populations

Ren Zhou*, Hongchen Zheng*, Mengfei Liu*, Zhen Liu*, Chuanhai Guo, Hongrui Tian, Fangfang Liu, Ying Liu, Yaqi Pan, Huanyu Chen, Zhe Hu, Hong Cai, Zhonghu He, Yang Ke

Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Laboratory of Genetics, Peking University Cancer Hospital & Institute, Beijing 100142, China

*These authors contributed equally to this work.

Correspondence to: Yang Ke. Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Laboratory of Genetics, Peking University Cancer Hospital & Institute, Beijing 100142, China. Email: keyang@bjmu.edu.cn; Zhonghu He. Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Laboratory of Genetics, Peking University Cancer Hospital & Institute, Beijing 100142, China. Email: zhonghuhe@foxmail.com.

Abstract

Objective: This study aimed to develop and validate a risk scoring system to identify high-risk individuals carrying malignant lesions in stomach for tailored gastric cancer screening.

Methods: A gastric cancer risk scoring system (GC-RSS) was developed based on questionnaire-based predictors for gastric cancer derived from systematic literature review. To assess the capability of this system for discrimination, risk scores for 8,214 and 7,235 outpatient subjects accepting endoscopic examination in two endoscopy centers, and 32,630 participants in a community-based cohort in China were calculated to plot receiver operating characteristic curves and generate area under the curve (AUC). To evaluate the performance of GC-RSS, the screening proportion, sensitivity and detection rate ratio compared to universal screening were used under different risk score cutoff values.

Results: GC-RSS comprised nine predictors including advanced age, male gender, low body mass index (<18.5 kg/m²), family history of gastric cancer, cigarette smoking, consumption of alcohol, preference for salty food, irregularity of meals and consumption of preserved food. This tool performed well in determining the risk of malignant gastric lesions with AUCs of 0.763, 0.706 and 0.696 in three validation sets. When subjects with risk scores ≥ 5 were evaluated with endoscopy, nearly 50% of these endoscopies could be saved with a detection rate of over 1.5 times achieved. When the cutoff was set at 8, only about 10% of subjects with the highest risk would be offered endoscopy, and detection rates for gastric cancer could be increased 2–4 fold compared to universal screening.

Conclusions: An effective questionnaire-based GC-RSS was developed and validated. This tool may play an important role in establishing a tailored screening strategy for gastric cancer in China.

Keywords: Cancer screening; external validation; gastric cancer; risk scoring system

Submitted Nov 02, 2021. Accepted for publication Nov 23, 2021.

doi: 10.21147/j.issn.1000-9604.2021.06.02

View this article at: <https://doi.org/10.21147/j.issn.1000-9604.2021.06.02>

Introduction

Gastric cancer is the fifth most commonly diagnosed cancer and the fourth leading cause of cancer-related deaths worldwide (1), and ranks even higher in China (2,3). It is estimated that incident cases of and deaths from gastric cancer in China constitute nearly half of cases and deaths worldwide (1). Due to the fact that patients are commonly diagnosed at an advanced stage (4), 5-year survival for gastric cancer in China was 35.1% according to cancer registry statistics during 2012–2015 (5), which is unsatisfactory compared to that of 60%–80% in Japan and South Korea (6,7). Fortunately, the 5-year survival for early-stage cancer patients can be as high as 90% (8), and early detection (screening) and treatment is thus a crucial strategy for the prevention and control of gastric cancer in China.

Endoscopic examination is currently the gold standard for screening for malignant gastric lesions (9), and it has been widely used in areas which are high-risk for upper gastrointestinal cancers since 2005 in China (10–12). Given the relatively low detection rates in population-based screening modalities (10–12) and the invasive nature of endoscopic examination and biopsy, identifying subjects who are at high risk for malignant gastric lesions without having recourse to large-scale endoscopic screening engenders concern (13).

In the Chinese experts consensus on screening of gastric cancer (2017 version), a risk scoring system was proposed to identify individuals at high risk for gastric cancer as the target for endoscopic screening (9). This consisted of five predictors, including age, gender, gastrin-17 (G-17), pepsinogen I/II ratio (PGR), and anti-*Helicobacter pylori* (*H. pylori*) antibody (9). Although this prediction tool was reported to be capable of identifying 69.6% of gastric cancers with a saving of 65.8% of endoscopic examinations in an external validation cohort (14), such inclusion of laboratory-based predictors would significantly increase the cost of risk-evaluation and impair generalizability when put into broader practice. In view of the huge population which needs screening in China, a non-invasive, easy-to-use, and accurate risk scoring system is needed as a pre-selection tool to identify individuals at high risk for malignant gastric lesions in order to determine whether further laboratory testing or endoscopic examination is warranted.

In this study, we established and evaluated a new questionnaire-based gastric cancer risk scoring system (GC-RSS) to identify individuals at high risk, based on two large-scale outpatient cohorts which received upper

gastrointestinal endoscopy, together with a community-based screening cohort to facilitate establishment of tailored gastric cancer screening in China.

Materials and methods

Establishment of GC-RSS

GC-RSS was based on questionnaire-based predictors reported in published studies regarding risk prediction for gastric cancer in Asian populations. We searched PubMed database for studies predicting the risk of gastric cancer. The following search algorithm with a combination of Medical Subject Headings terms and free text words was adopted to search the database: (gastric cancer OR stomach cancer OR stomach neoplasms) AND (prediction model OR risk assessment OR risk prediction OR risk scoring OR clinical decision rules). We first screened each study by titles, abstracts, and keywords, and then the full text of the studies meeting inclusion criteria were examined in detail. Eligible studies included in the further extraction of questionnaire-based candidate predictors were selected to meet the inclusion and exclusion criteria below. Inclusion criteria were: 1) risk prediction study or risk scoring study for gastric cancer; 2) sample size of >2,000; 3) study was conducted in Asian populations; and 4) study published after Jan 1st, 2010. The sole criterion for exclusion was any study that was a case-control study. We then extracted questionnaire-based candidate predictors from eligible studies, and predictors which met any one of the criteria below were candidate for further assessment: 1) biologically plausible for predicting risk of gastric cancer; or 2) a given predictor was reported in at least two independent risk prediction studies. Next, these candidate predictors were assessed in the three validation datasets using multivariable logistic regression. Predictors with an adjusted odds ratio over 1.2 in any validation dataset would be included in our risk scoring system. Finally, we assigned an integral score (weight) to each category of these predictors based on the effect sizes (odds ratios or hazard ratios) reported in published prediction models, and the sum of all predictors was defined as the final risk score for a specific individual.

Evaluation of GC-RSS

Study population

The performance of GC-RSS was evaluated in two outpatient cohorts which received upper gastrointestinal endoscopic examination (CCC2020010301, [© Chinese Journal of Cancer Research. All rights reserved.](http://chinaco-</p></div><div data-bbox=)

hort.bjmu.edu.cn) and in a community-based randomized controlled trial for endoscopic screening [Endoscopic Screening for Esophageal Cancer in China (ESECC) trial, NCT01688908]. These outpatient cohorts were enrolled from Southern and Northern China, and were designated the Southern outpatient cohort and the Northern outpatient cohort. The Southern outpatient cohort was established in June 2017 at the Peking University Shenzhen Hospital, which is one of the major hospitals in Shenzhen City, Guangdong Province, and the Northern outpatient cohort was established in March 2017 at the Hua County People's Hospital, which is a central hospital in Hua County, Henan Province (15). Participants in these two cohorts were recruited from outpatients in the endoscopy center who were undergoing upper gastrointestinal endoscopic examination with the following inclusion criteria: 1) aged 45–69 years; 2) no history of cancer, mental disorder, and no contraindications for endoscopy; and 3) completion of upper gastrointestinal endoscopy with a valid examination result. A total of 8,214 and 7,235 subjects who were recruited up to January 15th, 2021 in the Southern and Northern outpatient cohorts were enrolled in this study.

The ESECC trial was initiated in 2012 in Hua County, Anyang, Henan Province, which is a high-risk region for esophageal squamous cell carcinoma in China. A total of 668 villages in Hua County were randomly selected and allocated into the screening arm or control arm at a ratio of 1:1 by means of blocked randomization based on population size (16). The inclusion criteria were: 1) permanent residency in a target village; 2) aged 45–69 years with no history of endoscopic examination within 5 years of the initial interview; 3) no history of cancer, mental disorder and no contraindications for endoscopy; 4) patient negative for hepatitis B virus, hepatitis C virus and human immunodeficiency virus; and 5) patient agreed to complete all phases of this trial. After excluding esophageal cancer and cancer of the gastric cardia detected in the baseline endoscopic examination in the screening arm, 32,630 participants with sufficient questionnaire information in both arms were enrolled in this study.

This study was approved by the Institutional Review Board of the Peking University Cancer Hospital & Institute. Written informed consent was obtained from each subject in this study.

Data collection

Participants in two outpatient cohorts and the ESECC

cohort completed a standard computer-aided one-on-one questionnaire conducted by trained investigators and were then given a standard upper gastrointestinal endoscopy. Information about potential predictors for upper gastrointestinal cancer was collected, including demographic variables, lifestyle information, gastrointestinal-related symptoms, and family history of upper gastrointestinal cancer. During upper gastrointestinal endoscopy, the entire upper tract was carefully examined by experienced endoscopists, and biopsies of all focal lesions were taken. Biopsy specimens taken during endoscopy were fixed in 10% formaldehyde, embedded in paraffin, sectioned at 5 μ m and stained with hematoxylin and eosin (H&E). Pathologic diagnosis of biopsy specimens was performed by experienced pathologists. For the outpatient cohorts, the predicted outcome was defined as a diagnosis of malignant lesions in stomach including high-grade intraepithelial neoplasia (or severe dysplasia) of gastric mucosa, carcinoma *in situ*, and gastric cancer. For the ESECC cohort, malignant gastric lesions detected at index endoscopy and incident gastric cancers found in follow-up within 1 year after enrollment were defined as outcomes (17–19).

Statistical analysis

The demographic characteristics and life style variables among subjects enrolled in the Southern outpatient cohort, the Northern outpatient cohort, and the ESECC cohort were compared using the Chi-squared test for categorical variables and the Kruskal-Wallis rank-sum test for continuous variables.

To validate the performance of our GC-RSS in real-world modalities, we calculated the total risk score of each subject enrolled in this study based on our scoring system, and plotted the receiver operating characteristic (ROC) curves of GC-RSS separately for each of these three populations. The area under the curve (AUC) was also calculated for each population to quantify the capacity for discrimination of GC-RSS.

We also assumed a hypothetical tailored screening strategy wherein only subjects with risk scores above a specific cutoff were offered endoscopic examination. This was followed by calculation of the screening proportion, sensitivity, number of subjects requiring screening to identify one malignant gastric lesion, as well as the detection rate ratio compared to universal screening under different risk score cutoffs in these three populations.

All statistical analyses were conducted using R package

(Version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria). A two-sided $P < 0.05$ was considered to be statistically significant.

Results

Structure of GC-RSS

A total of five studies met both the inclusion and the exclusion criteria as determined by literature retrieval in the PubMed database. Predictors in prediction models established in previous studies are summarized in *Supplementary Table S1* (14,20-23). Based on the predictor selection criteria in this study, nine variables including advanced age, male gender, low body mass index (BMI) ($< 18.5 \text{ kg/m}^2$), family history of gastric cancer, cigarette smoking, consumption of alcohol, preference for salty food, irregularity of meals and consumption of preserved food were ultimately included in our risk scoring system (*Table 1*).

As shown in *Table 1*, age was categorized into four subgroups by 5-year intervals and assigned an integral score of 0–4. Gender was assigned as 0 for female or 2 for male, while other predictors were dichotomized with a score of 0 (No) or 1 (Yes). In this risk scoring system, the final score for each subject was calculated as the sum total of the scores for each predictor, which ranged from 0 to 13. A higher score reflected a higher risk of malignant gastric lesions in a given individual.

Performance of GC-RSS

The two outpatient cohorts receiving upper gastrointestinal endoscopic examination and the ESECC cohort were used to evaluate the performance of GC-RSS. As shown in *Table 2*, subjects in these three datasets show significant differences in demographic characteristics and life-style variables including age, gender, BMI, family history of gastric cancer, cigarette smoking, alcohol consumption and eating habits. The detection rates of malignant gastric lesions also varied significantly among these three datasets, which were 0.9% (78/8,214) in the Southern outpatient cohort, 0.4% (29/7,235) in the Northern outpatient cohort, and 0.1% (18/32,630) in the ESECC cohort.

When GC-RSS was applied to the three evaluation datasets, the median of the risk score was 4 in the Southern and Northern outpatient cohorts, and 5 in the ESECC cohort. When considering the pathologic diagnosis at endoscopy for each individual, subjects with higher scores

Table 1 Variables included in GC-RSS and their weights

Variables*	Scores
Age (year)	
45–49	0
50–54	1
55–59	2
60–64	3
65–69	4
Gender	
Female	0
Male	2
Family history of gastric cancer**	
No	0
Yes	1
Cigarette smoking#	
No	0
Yes	1
Alcohol consumption##	
No	0
Yes	1
Body mass index (kg/m ²)	
≥ 18.5	0
< 18.5	1
Consumption of salty food (time per week)	
< 1	0
≥ 1	1
Meal irregularity	
No	0
Yes	1
Consumption of preserved food (time per week)	
< 1	0
≥ 1	1

*, variables were questionnaire-based predictors which met any one of the criteria below: 1) biologic plausibility in predicting the risk of gastric cancer or 2) reported in at least two published risk prediction studies; **, family history of gastric cancer was defined as gastric cancer cases in the immediate family and relatives within three generations; #, cigarette smoking was defined as a smoking history of at least 18 packs per year; ##, alcohol drinking was defined as drinking Chinese liquor (containing $> 40\%$ alcohol) at least twice a week for ≥ 12 months. GC-RSS, gastric cancer risk scoring system.

showed higher detection rates for malignant gastric lesions. When scores were ≤ 8 , detection rates were below 2% in the Southern outpatient cohort and below 0.2% in the ESECC cohort, while rates rose separately to nearly 10.0%

Table 2 Selected demographic characteristics and life-style variables among individuals in Southern outpatient cohort, Northern outpatient cohort and ESECC cohort

Variables*	n (%)			P**
	Southern outpatient cohort (N=8,214)	Northern outpatient cohort (N=7,235)	ESECC cohort (N=32,630)	
Age (year)				<0.001
45–49	1,870 (22.8)	1,563 (21.6)	7,119 (21.8)	
50–54	2,172 (26.4)	2,015 (27.9)	6,549 (20.1)	
55–59	1,836 (22.4)	1,336 (18.5)	6,334 (19.4)	
60–64	1,425 (17.3)	1,229 (17.0)	7,695 (23.6)	
65–69	911 (11.1)	1,092 (15.1)	4,933 (15.1)	
Gender				<0.001
Female	4,510 (54.9)	4,244 (58.7)	16,706 (51.2)	
Male	3,704 (45.1)	2,991 (41.3)	15,924 (48.8)	
Family history of gastric cancer				<0.001
No	7,812 (95.1)	6,870 (95.0)	31,777 (97.4)	
Yes	402 (4.9)	365 (5.0)	853 (2.6)	
Cigarette smoking				<0.001
No	5,947 (72.4)	5,390 (74.5)	21,397 (65.6)	
Yes	2,267 (27.6)	1,845 (25.5)	11,233 (34.4)	
Alcohol consumption				<0.001
No	6,898 (84.0)	5,921 (81.8)	24,749 (75.8)	
Yes	1,316 (16.0)	1,314 (18.2)	7,881 (24.2)	
Body mass index (kg/m ²)				<0.001
≥18.5	7,794 (94.9)	7,113 (98.3)	32,126 (98.5)	
<18.5	420 (5.1)	122 (1.7)	504 (1.5)	
Consumption of salty food (time per week)				<0.001
<1	3,211 (39.1)	2,381 (32.9)	6,999 (21.4)	
≥1	5,003 (60.9)	4,854 (67.1)	25,631 (78.6)	
Meal irregularity				<0.001
No	6,754 (82.2)	6,369 (88.0)	29,711 (91.1)	
Yes	1,460 (17.8)	866 (12.0)	2,919 (8.9)	
Consumption of preserved food (time per week)				<0.001
<1	4,017 (48.9)	4,708 (65.1)	21,233 (65.1)	
≥1	4,197 (51.1)	2,527 (34.9)	11,397 (34.9)	
Malignant gastric lesions				<0.001
No	8,136 (99.1)	7,206 (99.6)	32,612 (99.9)	
Yes	78 (0.9)	29 (0.4)	18 (0.1)	

ESECC, endoscopic screening for esophageal cancer in China; *, Variables were risk factors associated with gastric cancer from biologic aspect or predictors reported in as least two previous studies; **, P values were obtained from the Chi-squared test for categorical variables.

and 0.4% in these two cohorts when scores increased to ≥ 10 (*Supplementary Figure S1*). ROC curves demonstrated

that the risk scoring system had a satisfactory performance in discriminating individuals at high risk for malignant

gastric lesions both in the outpatient endoscopic screening cohorts, and in the community-based ESECC cohort (Figure 1), with an AUC of 0.763 [95% confidence interval (95% CI): 0.707–0.819] for the Southern outpatient cohort, 0.706 (95% CI: 0.606–0.807) for the Northern outpatient cohort, and 0.696 (95% CI: 0.567–0.825) for the ESECC cohort, respectively.

We also assessed the effectiveness of the risk scoring system by assuming a hypothetical tailored screening wherein only subjects with a risk score above a certain cutoff were offered endoscopy. When the cutoff for risk score was set as 5 to ensure a population coverage of about 50% (as required in the current national screening programs in China), detection rates of 1.8 times, 1.5 times, and 1.6 times as compared to universal screening could be achieved in the Southern and the Northern outpatient cohorts as well as in the ESECC cohort, respectively (Table 3). If a more stringent cutoff of 8 was chosen in order to enrich for patients with malignant gastric lesions under a given workload, only 10.9%, 9.9% and 14.1% of subjects would need to accept endoscopic examination, with detection rate ratios of 4.0, 3.8 and 2.4 (Table 3) as compared to universal examination in these three cohorts.

Discussion

Taking the huge population size of China into consideration, together with the high cost of endoscopy and potential harm of this kind of invasive examination, nationwide universal screening for gastric cancer is impractical in China. Identifying individuals at high risk for

malignant gastric lesions in the general population is therefore a critical antecedent to selecting groups for large-scale endoscopic screening. In this study, we constructed and validated a simple questionnaire-based risk scoring system for malignant gastric lesions, and application of this tool would markedly facilitate pre-selection of targets for hospital-based and community-based endoscopic screening.

Our GC-RSS was of good rationality and showed ideal performance in terms of discrimination. This was validated in two independent outpatient screening cohorts and a community-based screening cohort. Although these three populations were heterogeneous in terms of demographic characteristics (e.g., age and gender), economic development level, personal lifestyle (e.g., cigarette smoking, alcohol consumption and eating habits) and the prevalence of malignant gastric lesions, the GC-RSS showed robust performance with AUCs of 0.763, 0.706 and 0.696, respectively, which demonstrated its good generalizability and potential for practical application in gastric cancer screening modalities in varied populations.

Among the nine variables included in our GC-RSS, advanced age, male gender, family history of gastric cancer, cigarette smoking, alcohol consumption and preference for salty food are widely accepted risk factors for malignant gastric lesions (24–26). Other predictors such as low BMI and ingestion of preserved food have also previously been reported to be associated with a higher risk of malignant gastric lesions (27,28). Since age was reported to have the strongest effect among these nine predictors (14), we assigned weight of 0–4 to the age by 5-year intervals in our risk scoring system. Gender was also an important

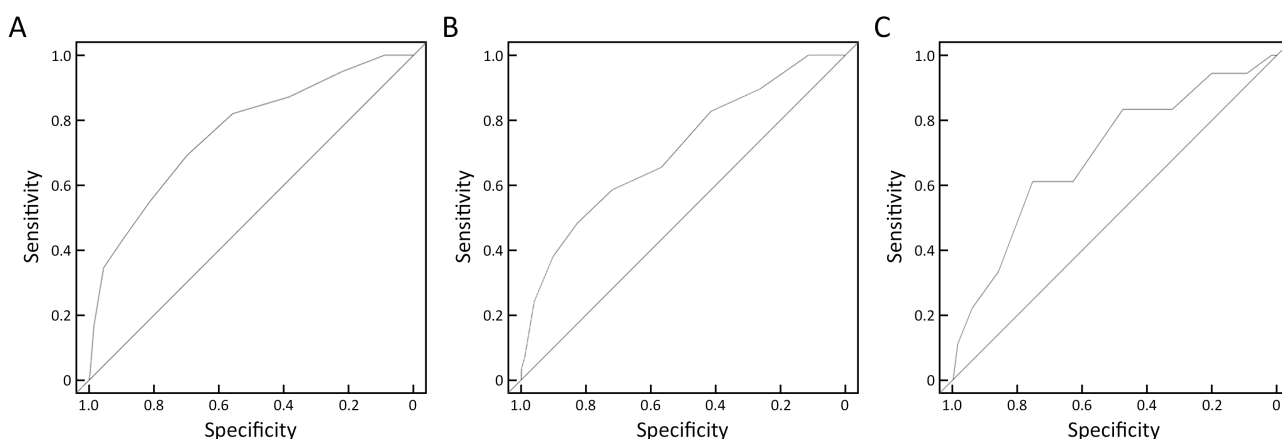


Figure 1 Receiver operating characteristic curves of the risk scoring system for gastric cancer in (A) Southern outpatient cohort (AUC=0.763); (B) Northern outpatient cohort (AUC=0.706); and (C) ESECC cohort (AUC=0.696). AUC, area under the curve; ESECC, endoscopic screening for esophageal cancer in China.

Table 3 Effectiveness of GC-RSS in Southern outpatient cohort, Northern outpatient cohort, and ESECC cohort when risk score cutoff was set from 0 to 10

Cutoff	Southern outpatient cohort (N=8,214)				Northern outpatient cohort (N=7,235)				ESECC cohort (N=32,630)			
	Screening proportion (Sensitivity) (%)	No. of endoscopy to detect per case	DR (%)	DRR (95% CI) compared to universal screening (%)	Screening proportion (Sensitivity) (%)	No. of endoscopy to detect per case	DR (%)	DRR (95% CI) compared to universal screening (%)	Screening proportion (Sensitivity) (%)	No. of endoscopy to detect per case	DR (%)	DRR (95% CI) compared to universal screening (%)
10	1.6 (16.7)	10.3	9.7	10.2 (5.1–19.1)	1.1 (6.9)	40.5	2.5	6.2 (0.7–25.0)	1.6 (11.1)	256.0	0.4	7.1 (0.8–29.7)
9	4.8 (34.6)	14.6	6.9	7.2 (4.4–11.5)	4.1 (24.1)	42.4	2.4	5.9 (2.2–13.8)	6.1 (22.2)	495.8	0.2	3.7 (0.9–11.1)
8	10.9 (43.6)	26.4	3.8	4.0 (2.6–6.1)	9.9 (37.9)	65.2	1.5	3.8 (1.7–7.9)	14.1 (33.3)	768.0	0.1	2.4 (0.8–6.2)
7	19.2 (55.1)	36.7	2.7	2.9 (1.9–4.2)	17.4 (48.3)	90.0	1.1	2.8 (1.3–5.4)	24.7 (61.1)	733.8	0.1	2.5 (1.1–5.5)
6	30.6 (69.2)	46.6	2.2	2.3 (1.6–3.3)	28.3 (58.6)	120.4	0.8	2.1 (1.1–3.9)	37.2 (61.1)	1,102.5	0.1	1.6 (0.7–3.7)
5	44.7 (82.1)	57.3	1.7	1.8 (1.3–2.6)	43.4 (65.5)	165.1	0.6	1.5 (0.8–2.8)	52.5 (83.3)	1,142.5	0.1	1.6 (0.7–3.3)
4	61.9 (87.2)	74.8	1.3	1.4 (1.0–2.0)	58.7 (82.8)	176.8	0.6	1.4 (0.8–2.5)	67.7 (83.3)	1,473.0	0.1	1.2 (0.6–2.6)
3	77.9 (94.9)	86.5	1.2	1.2 (0.9–1.7)	73.8 (89.7)	205.2	0.5	1.2 (0.7–2.1)	79.9 (94.4)	1,533.3	0.1	1.2 (0.6–2.4)
2	91.2 (100)	96.0	1.0	1.1 (0.8–1.5)	88.7 (100)	221.0	0.5	1.1 (0.6–2.0)	90.7 (94.4)	1,741.8	0.1	1.0 (0.5–2.1)
1	97.9 (100)	103.1	1.0	1.0 (0.7–1.4)	97.8 (100)	244.0	0.4	1.0 (0.6–1.8)	98.3 (100)	1,782.7	0.1	1.0 (0.5–2.1)
0	100 (100)	105.3	1.0	1.0 (0.7–1.4)	100 (100)	249.3	0.4	1.0 (0.6–1.7)	100 (100)	1,812.8	0.1	1.0 (0.5–2.0)

GC-RSS, gastric cancer risk scoring system; ESECC, endoscopic screening for esophageal cancer in China; DR, detection rate; DRR, detection rate ratio; 95% CI, 95% confidence interval.

predictor for risk of malignant gastric lesions with an effect size of 2–3, which was slightly lower than that of age (14,20,21); thus a weight of 0 was assigned for female gender and 2 for male gender. For other predictors, the effect sizes (odds ratios or hazard ratios) were reported to be slightly higher than 1 (14,20,21,23), and we thus assigned the same weight of 0 or 1 to these predictors as we thought there were no essential differences among them.

For application of the GC-RSS, there are two main scenarios where it may play a role in gastric cancer screening. First, GC-RSS can be utilized to select the target for endoscopic screening in large-scale community-based screening programs which usually are designed for population coverage and sensitivity. When one accepts the requirements of national early detection and early treatment projects in China, and intends to initiate community-based screening under a population coverage of about 50% (29), a cutoff of 5 can be chosen to select subjects to undergo endoscopy. In this case, the detection rate ratio can be increased by 50%–80% as compared to current practice, and the number of endoscopies required to detect one case can be reduced from 105 to 57 in the Southern outpatient cohort, and from 249 to 165 in the Northern outpatient cohort. Second, when resources are extremely limited, GC-RSS may also show excellent capacity of enrichment for patients with malignant gastric lesions which is thus suitable for hospital-based screening,

as it relies more on the enrichment of cancer cases and is an important gastric cancer screening modality (30). For example, if only about 10% of subjects under evaluation by questionnaire investigation can be endoscopically examined in a resource-limited setting, a cutoff of 8 for our risk scoring system can increase the detection rate 2–4 times as compared to universal screening. At the same time, it may result in reduction of the number of endoscopies to detect one case from 105 to 26 in the Southern outpatient cohort, and from 249 to 65 in the Northern outpatient cohort.

In a previous risk evaluation tool proposed by the Chinese experts' consensus in 2014, individuals of ≥40 years old were evaluated based on the following criteria: 1) living in areas of high risk for gastric cancer; 2) history of *H. pylori* infection; 3) having precancerous diseases for gastric cancer (e.g. chronic atrophic gastritis and gastric ulcer); 4) family history of gastric cancer among first-degree relatives; and 5) other positive risk factors for gastric cancer (e.g. intake of salty food, intake of preserved food, cigarette smoking and alcohol consumption), and subjects were defined as high risk once they met any one of these five criteria. This rough tool would allow a high proportion of subjects under evaluation to accept endoscopic screening (e.g. all residents in high-risk areas), and thus have no practical meaning in real-world gastric cancer screening efforts. In a hospital-based endoscopic screening program involving 115 centers in China (14), this tool resulted in an

unimproved detection rate for gastric cancer (4). Another questionnaire-based risk scoring system established by Cao *et al.* was reported to have a lower sensitivity of 62% in identification of high-risk individuals for gastric cancer with a screening coverage of nearly 50% in high-risk areas for esophageal cancer (22). Moreover, this scoring system, which was established with an esophageal cancer specific cohort, to some extent lacked rationality for risk prediction of gastric cancer as it consisted of a series of esophageal cancer related predictors. In addition, its performance in predicting gastric cancer was not validated in external populations, and as such the real-world performance of this scoring system is largely undetermined.

This study has several strengths. First, our risk scoring system was based on questionnaire data, which greatly improved the feasibility of its application in real-world screening efforts. Second, the performance of this risk scoring system was validated in two clinical outpatient cohorts and a large-scale community-based population with heterogeneous characteristics. Third, this newly proposed risk scoring system was established specially for malignant gastric lesions based on predictors with sufficient biologic plausibility for malignancy in the stomach.

This study also has limitations. First, the absolute number of patients with malignant gastric lesions in the three cohorts that we used was relatively limited. Thus, further evaluation is still needed, especially in areas of high risk for gastric cancer with larger numbers of cancer cases. In addition, simple integers were assigned when determining the weight of each predictor in this study, and determination of more precise weighting may be a direction for updating our risk scoring system when larger population and additional cases are available in the future.

Conclusions

We have provided and validated a questionnaire-based, easy-to-use, low-cost and accurate GC-RSS to identify high-risk individuals for malignant gastric lesions in Chinese populations. This tool may play an important role in the establishment of a tailored screening strategy for gastric cancer, in both China and other areas with situations that are similar in terms of gastric cancer etiology and epidemiology.

Acknowledgements

This study was supported by the National Science &

Technology Fundamental Resources Investigation Program of China (No. 2019FY101102); the National Natural Science Foundation of China (No. 82073626, 81773501); the National Key R&D Program of China (No. 2016YFC0901404); the Beijing-Tianjin-Hebei Basic Research Cooperation Project (No. J200016); the Digestive Medical Coordinated Development Center of Beijing Hospitals Authority (No. XXZ0204); the Beijing Hospitals Authority Youth Programme (No. QML20201101); Sanming Project of Shenzhen (No. SZSM201612061); and the Beijing Nova Program (No. Z201100006820093).

Footnote

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

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209-49.
2. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016;66:115-32.
3. Zhang S, Sun K, Zheng R, et al. Cancer incidence and mortality in China, 2015. *J Natl Cancer Center* 2021;1:2-11.
4. Chinese Society of Digestive Endoscopy, Chinese Anti-Cancer Association. Experts' consensus on the endoscopic screening, diagnosis and treatment of early gastric cancer in China (2014 Changsha). *Zhonghua Xiao Hua Za Zhi* (in Chinese) 2014;34:433-48.
5. Zeng H, Chen W, Zheng R, et al. Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. *Lancet Glob Health* 2018;6:e555-67.
6. Ito Y, Miyashiro I, Ito H, et al. Long-term survival and conditional survival of cancer patients in Japan using population-based cancer registry data. *Cancer Sci* 2014;105:1480-6.
7. Hong S, Won YJ, Lee JJ, et al. Cancer statistics in Korea: Incidence, mortality, survival, and prevalence in 2018. *Cancer Res Treat* 2021;53:301-15.
8. Pyo JH, Lee H, Min BH, et al. Long-term outcome of endoscopic resection vs. surgery for early gastric

- cancer: A non-inferiority-matched cohort study. *Am J Gastroenterol* 2016;111:240-9.
9. National Clinical Research Center for Digestive Diseases, Chinese Society of Digestive Endoscopy, Health Management Association of Chinese Medical Association. China experts' consensus on the protocol of early gastric cancer screening (2017, Shanghai). *Wei Chang Bing Xue* (in Chinese) 2018;23:92-7.
 10. Guo L, Zhang S, Liu S, et al. Determinants of participation and detection rate of upper gastrointestinal cancer from population-based screening program in China. *Cancer Med* 2019;8:7098-107.
 11. Zheng X, Mao X, Xu K, et al. Massive endoscopic screening for esophageal and gastric cancers in a high-risk area of China. *PLoS One* 2015;10:e0145097.
 12. Zeng H, Sun K, Cao M, et al. Initial results from a multi-center population-based cluster randomized trial of esophageal and gastric cancer screening in China. *BMC Gastroenterol* 2020;20:398.
 13. He Z, Ke Y. Precision screening for esophageal squamous cell carcinoma in China. *Chin J Cancer Res* 2020;32:673-82.
 14. Cai Q, Zhu C, Yuan Y, et al. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population: a nationwide multicentre study. *Gut* 2019;68:1576-87.
 15. Liu Z, Guo C, He Y, et al. A clinical model predicting the risk of esophageal high-grade lesions in opportunistic screening: a multicenter real-world study in China. *Gastrointest Endosc* 2020;91:1253-60.e3.
 16. He Z, Liu Z, Liu M, et al. Efficacy of endoscopic screening for esophageal cancer in China (ESECC): design and preliminary results of a population-based randomised controlled trial. *Gut* 2019;68:198-206.
 17. Shi C, Liu M, Liu Z, et al. Using health insurance reimbursement data to identify incident cancer cases. *J Clin Epidemiol* 2019;114:141-9.
 18. Tian H, Xu R, Li F, et al. Identification of cancer patients using claims data from health insurance systems: A real-world comparative study. *Chin J Cancer Res* 2019;31:699-706.
 19. Tian H, Yang W, Hu Y, et al. Estimating cancer incidence based on claims data from medical insurance systems in two areas lacking cancer registries in China. *EClinicalMedicine* 2020; 20:100312.
 20. Iida M, Ikeda F, Hata J, et al. Development and validation of a risk assessment tool for gastric cancer in a general Japanese population. *Gastric Cancer* 2018;21:383-90.
 21. Charvat H, Sasazuki S, Inoue M, et al. Prediction of the 10-year probability of gastric cancer occurrence in the Japanese population: the JPHC study cohort II. *Int J Cancer* 2016;138:320-31.
 22. Cao M, Li H, Sun D, et al. Classifying risk level of gastric cancer: Evaluation of questionnaire-based prediction model. *Chin J Cancer Res* 2020;32:605-13.
 23. Eom BW, Joo J, Kim S, et al. Prediction model for gastric cancer incidence in Korean population. *PLoS One* 2015;10:e0132613.
 24. 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.
 25. Ma K, Baloch Z, He TT, et al. Alcohol consumption and gastric cancer risk: A meta-analysis. *Med Sci Monit* 2017;23:238-46.
 26. den Hoed CM, Kuipers EJ. Gastric cancer: How can we reduce the incidence of this disease? *Curr Gastroenterol Rep* 2016;18:34.
 27. Bae JM. Body mass index and risk of gastric cancer in Asian adults: A meta-epidemiological meta-analysis of population-based cohort studies. *Cancer Res Treat* 2020;52:369-73.
 28. Ren JS, Kamangar F, Forman D, et al. Pickled food and risk of gastric cancer-a systematic review and meta-analysis of English and Chinese literature. *Cancer Epidemiol Biomarkers Prev* 2012;21:905-15.
 29. Wei W, Hao C, Guan C, et al. Esophageal histological precursor lesions and subsequent 8.5-year cancer risk in a population-based prospective study in China. *Am J Gastroenterol* 2020;115:1036-44.
 30. Wang G, Wei W. A new transition of the screening, early diagnosis and early treatment project of the upper gastrointestinal cancer: opportunistic screening. *Zhonghua Yu Fang Yi Xue Za Zhi* (in Chinese) 2019;53:1084-7.

Cite this article as: Zhou R, Zheng H, Liu M, Liu Z, Guo C, Tian H, Liu F, Liu Y, Pan Y, Chen H, Hu Z, Cai H, He Z, Ke Y. Development and validation of a questionnaire-based risk scoring system to identify individuals at high risk for gastric cancer in Chinese populations. *Chin J Cancer Res* 2021;33(6):649-658. doi: 10.21147/j.issn.1000-9604.2021.06.02

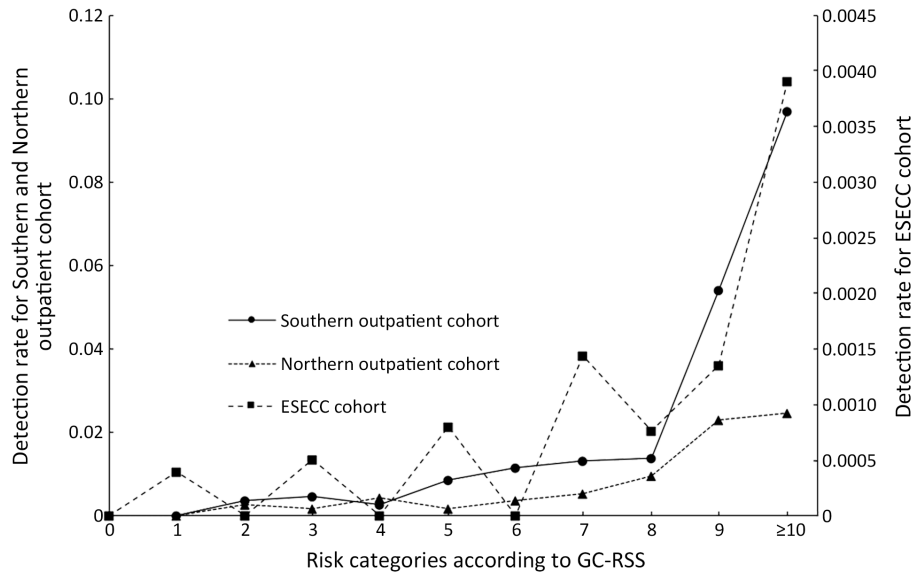


Figure S1 Detection rates of assumptive tailored screening with different cutoffs when GC-RSS was applied in Southern outpatient cohort, Northern outpatient cohort, and ESECC cohort. GC-RSS, gastric cancer risk scoring system; ESECC, endoscopic screening for esophageal cancer in China.

Table S1 Predictors in previously published studies regarding the risk prediction of gastric cancer among Asian populations

Study	Country	Study design	Predictors
Eom <i>et al.</i> 2015 (23)	Korea	Prospective study	Age, BMI, family history, meal regularity, salt preference, drinking, smoking and physical activity
Charvat <i>et al.</i> 2016 (21)	Japan	Prospective study	Age, gender, age × gender, smoking, highly salted food consumption, family history of gastric cancer, anti-HP IgG and serum pepsinogen statuses
Iida <i>et al.</i> 2018 (20)	Japan	Prospective study	Age, sex, the combination of <i>H. pylori</i> antibody and pepsinogen status, hemoglobin A1c level, and smoking
Cai <i>et al.</i> 2019 (14)	China	Cross-section study	Age, sex, PG I/II, G-17 level, <i>H. pylori</i> infection, pickled food and fried food
Cao <i>et al.</i> 2020 (22)	China	Prospective study	Smoking, drinking, eating salt-preserved food, eating very hot and hard food, family history of upper gastrointestinal cancer, current symptom of chest pain, pressure or burning, dysplasia, chronic heartburn or indigestion, vomiting or hemoptysis, progressive weight loss, esophageal reflux, and history of gastric disease

BMI, body mass index; *H. pylori*, *Helicobacter pylori*; IgG, immunoglobulin G; PG, pepsinogen; G-17, gastrin-17.