

A nomogram for endoscopic screening in a high esophageal squamous cell cancer risk area: results from a population-based study

Jie Xing,¹ Li Min,¹ Hao Zhang,² Peng Li,¹ Wei Li,¹ Fujin Lv,¹ Yongjun Wang,¹ Zheng Zhang,¹ Hengcun Li,¹ Qingdong Guo,¹ Siyi Wang,¹ Yu Zhao,¹ Junmin Wang,³ Xiaoyan Shi,⁴ Anxin Wang,⁵ Shengtao Zhu,¹ Ming Ji,¹ Yongdong Wu,¹ Shutian Zhang¹

On behalf of the NCECS Study Group

¹Department 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, People's Republic of China; ²Department of Gastroenterology, Jizhong Energy Fengfeng Group Hospital, Handan, Hebei 056200, People's Republic of China; ³Department of Gastroenterology, Handan Central Hospital, Handan, Hebei 056001, People's Republic of China; ⁴Department of Pathology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, People's Republic of China; ⁵Department of Epidemiology, Beijing Tiantan Hospital, Capital Medical University, Beijing 100029, People's Republic of China

Correspondence: Shutian Zhang
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, 95 Yong'an Road, Beijing 100050, People's Republic of China
Email zhangshutian@ccmu.edu.cn

Background: Endoscopy is the main approach used for esophageal squamous cell carcinoma (ESCC) screening, especially in high-risk areas. However, little consensus has been achieved in recent ESCC screening programs, and endoscopists have selected patients only by age and family history.

Patients and methods: To generate a proper strategy for selecting an eligible population for endoscopic screening based on demographic factors, lifestyle, and eating habits, a total of 7,830 residents in an area with a high risk of ESCC were recruited for this study. All participants underwent endoscopic examinations that were conducted by experienced endoscopists. Risk factors for ESCC and other lesions were selected by univariate and multivariate logistic regressions. A nomogram for the prediction of ESCC and premalignant lesions was constructed, which included information on age, sex, occupation, labor intensity, income, and mining exposure. Receiver operating characteristic (ROC) curve analysis was performed to present the predictive accuracy of the nomograms.

Results: The area under the curve (95% CI) was 0.749 (0.711–0.788) for this nomogram. By applying this nomogram, we could exclude 60% (4704/7830) of patients before endoscopy screening and detect all ESCC cases as well as most esophageal lesions in the remaining population.

Conclusion: In conclusion, we provided a ready-to-use preclinical tool with the potential to select eligible people with high risk of ESCC for endoscopy screening.

Keywords: nomogram, endoscopy screening, ESCC, population-based study

Introduction

With a total of 4,56,000 new cases and nearly 4,00,000 deaths in 2012, esophageal cancer (EC) ranks as the eighth most common malignant carcinoma and the sixth leading cause of death worldwide.¹ Esophageal adenocarcinoma (EAC) is the main subtype of EC in western countries, and its incidence has increased dramatically in the past decades.² However, esophageal squamous cell carcinoma (ESCC) remains the dominant subtype in East Asia.³ Though ESCC has steadily decreased along with various socioeconomic developments and changes in lifestyle, it remains the greatest cancer burden in some high-risk areas.³

Due to deficiencies in the serous layer, ESCC is highly aggressive and easily invades neighboring tissues.⁴ Most ESCC patients are diagnosed at a late stage and have a very poor prognosis.⁵ Kumar et al reported that the overall 5-year survival of ESCC is only 9%, but with early detection and treatment, the survival rate increases significantly.⁶ Morimoto et al from Japan reported that the 5-year survival rates decreased dramatically if there had been a failure to diagnose the disease at a very early stage, ie, the 5-year survival rates for stage 0 ESCC was 83% but only 47% for

stage I ESCC.⁷ Increasing the early detection rate of ESCC would be very beneficial for reducing the medical burden of high ESCC risk areas.⁸

Recently, for EAC, premalignant and precursor lesions have been well studied. Risk factors such as esophagitis, gastroesophageal reflux diseases, Barrett esophagus, drinking, and smoking are widely recognized and discussed in EAC screening guidelines.⁹ However, for ESCC, many risk factors have been proposed, such as age, sex, smoking, and drinking, but little consensus has been reached.^{10,11} Due to the lack of clinical and epidemiological evidence to select eligible people with high risk, ESCC screening has been nearly totally dependent on endoscopy.⁹ Recently, ESCC screening programs selected patients only by age and family history,¹²⁻¹⁴ which resulted in low compliance in screening and wasted medical resources.

In this study, we established a cohort undergoing endoscopic screening in Handan City, Hebei province (located in the Taihang Mountain area, a famous high ESCC risk area in People's Republic of China), to investigate the independent risk factors of ESCC and to establish a method to select the eligible

population for endoscopic screening based on demographic data, lifestyle factors, eating habits, and psychological factors.

Patients and methods

Study population and inclusion/exclusion criteria

The study was approved by Medical Ethics Committee of Beijing Friendship Hospital, Capital Medical University in compliance with the Declaration of Helsinki. A total of 10,135 residents between 40 and 70 years old in Handan City, Hebei Province (where the incidence of ESCC is higher than in other regions of People's Republic of China) were recruited during October 2010 to September 2011 for this study. Written informed consent was obtained from 8,040 residents, and 8,027 finished the epidemiology questionnaire. General health examinations were performed, and 197 people with iodine allergy, any other tumors, hyperthyreosis, pregnancy, lactation, and psychosis were excluded. After questionnaire administration and physical examination, a total of 7,830 participants were enrolled and successfully completed an upper gastrointestinal endoscopy examination (Figure 1).

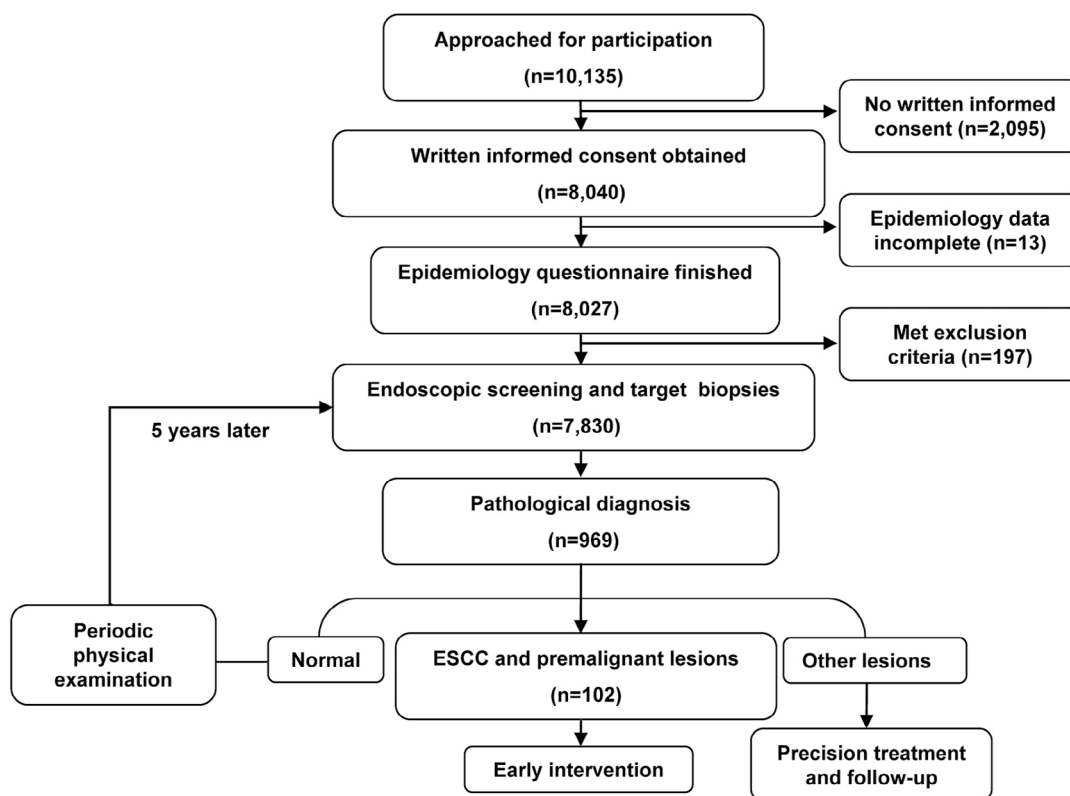


Figure 1 Flowchart of participant inclusion and exclusion.

Notes: A total of 10,135 residents were recruited in this study. Of these, 8,040 residents provided written informed consent and 8,027 completed the epidemiology questionnaire. Then, 197 people were excluded due to the exclusion criteria (iodine allergy, hyperthyreosis, pregnancy, etc.). Finally, 7,830 participants were enrolled and successfully completed an upper gastrointestinal endoscopy examination.

Endoscopy performance and quality control

All of the endoscopic examinations and therapies were completed by experienced endoscopists of the National Clinical Research Center for Digestive Disease. Fasting for at least 8 hours was required, and anticoagulant drugs were to be stopped 1 week before endoscopy. Iodine staining was applied to facilitate the discovery of lesions. Any esophageal lesions including carcinoma, dysplasia, hyperplasia, inflammatory lesions, ulcers, and reflux esophagitis were biopsied, and 5% of normal esophageal mucosa were also examined as negative controls. Each specimen of biopsied tissue was examined by two pathologists using the World Health Organization 2010 digestive oncopathologic diagnostic criteria (representative pathological and endoscopic images are displayed in Figure 2).

Data acquisition and processing

Questionnaires were mailed to participants to collect data regarding demographics, lifestyle factors, eating habits,

psychological factors, etc. Smoking status was divided into four subgroups: yes (smoking >6 months and ≥ 1 cigarettes per day), occasionally (smoking <1 cigarette/day or <6 months), refrained (ie, stopped smoking for >6 months), and never (never smoked). Eating habits (including leek food intake, fresh vegetable consumption, etc.) was defined by two subgroups: yes (≥ 3 times per month) and rarely or never (never or <3 times per month). Endoscopic examinations were performed to collect data on different esophageal diseases including reflux esophagitis, dysplasia, ulcers, etc. The data regarding pathological diagnoses were based on the consensus of two pathologists. All questionnaires and clinical information were entered into the database “back to back” by two researchers independently.

Statistical analyses

All statistical analyses were performed using R Software 3.2.2 (www.r-project.org). Packages “plyr”, “dplyr”, and “reshape2” were loaded for data cleaning, and package “rms”

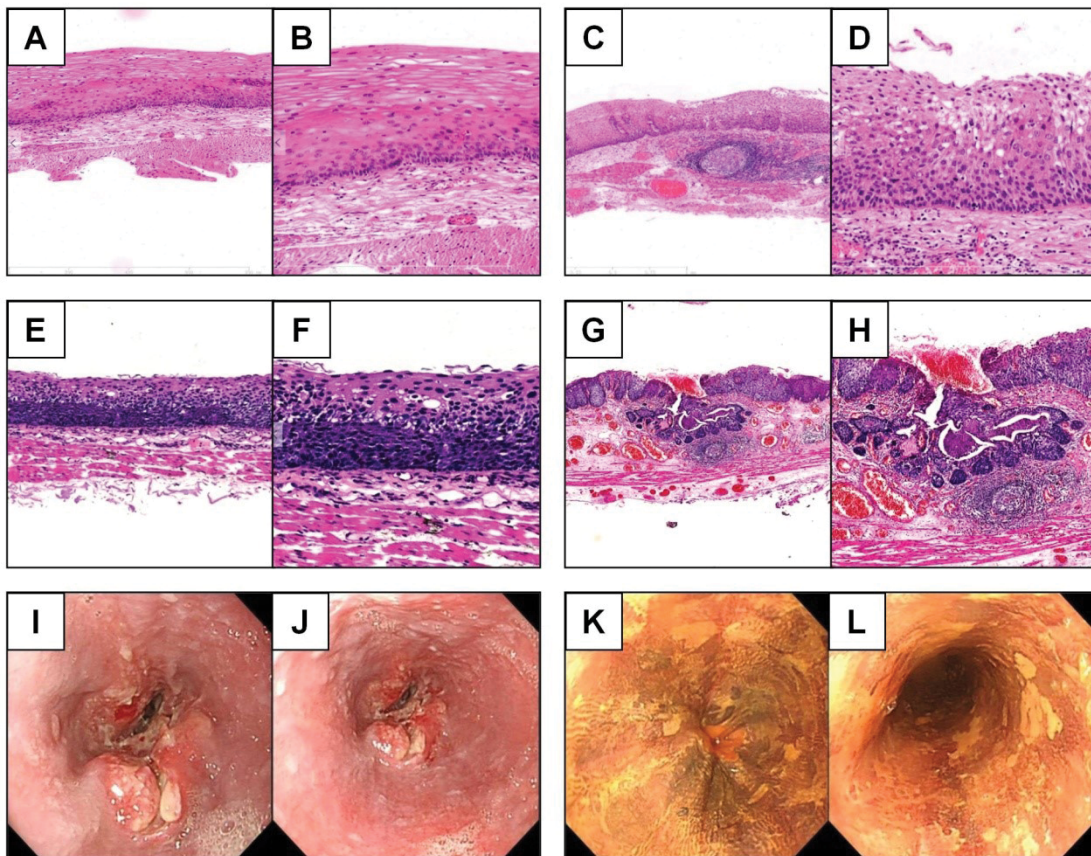


Figure 2 Representative immunological and endoscopic images of different esophageal diseases.

Notes: Representative immunohistochemical images of esophageal low-grade dysplasia (A. 20 \times , B. 40 \times), esophageal middle-grade dysplasia (C. 20 \times , D. 40 \times), esophageal high-grade dysplasia (E. 20 \times , F. 40 \times), and ESCC (G. 20 \times , H. 40 \times). Representative endoscopic images of uneven surface of the esophagus (I, J) and an iodine-unstained area in the esophageal mucosa (K, L).

Abbreviation: ESCC, esophageal squamous cell carcinoma.

was used for nomogram building. Univariate analysis was performed to identify risk factors for esophageal lesions, and an OR with 95% CI of potential risk factors was displayed. Then, multivariate regression analysis was used to select independent influence factors, and nomograms were built based mainly on these results. A receiver operating characteristic (ROC) curve was used to present the predictive accuracy of nomograms graphically. The area under the curve (AUC) for validation was applied to evaluate the accuracy of the nomograms. We performed calibration for the established nomograms and applied 1,000 repetitions of bootstrap sample corrections to internally validate the nomograms. All statistical results with $P < 0.05$ were regarded as statistically significant.

Results

Descriptive analysis of the patient population

Overall, 7,830 patients successfully completed an upper gastrointestinal endoscopy examination, and detailed data on demographics, lifestyle factors, eating habits, occupational factors, and disease history were collected. The most common esophageal lesions represented in this population were reflux esophagitis (15.22%, $n=1,193$), inflammatory lesions (5.68%, $n=445$), and dysplasia (1.25%, $n=98$). Only four carcinomas were detected (0.05%), and all of them were ESCC. There were 14 cases of hyperplasia (0.18%); 10 cases of esophageal ulcer (0.13%) were also represented (Table 1).

Based on the consensus of the participating local physicians and endoscopists of National Clinical Research Center for Digestive Disease, it was suggested that given the available data (low detection rates of carcinoma and dysplasia), a nomogram to predict carcinoma and premalignant lesions (dysplasia) risk would be very helpful in shrinking the population eligible for endoscopy screening.

Detect risk factors for ESCC and premalignant lesions

To generate a nomogram for the prediction of ESCC and premalignant lesions, we summarized the univariate logistic regression analyses of ESCC/premalignant lesions (Table 2 and Table S1). Age, sex, education, occupation, labor intensity, income, mining exposure, hypertension, and tea consumption were all associated with ESCC and premalignant lesions risk (all $P < 0.05$). However, in the multivariate analysis, only age, sex, and occupation were

Table 1 Esophageal lesions in 7,830 cases

Lesions	Number	Percentage
Carcinoma	4	0.05
Dysplasia	98	1.25
Hyperplasia	14	0.18
Inflammatory lesions	445	5.68
Esophageal ulcer	10	0.13
Reflux esophagitis	1,193	15.24
Normal esophagus	6,066	77.47

independent predictive factors for ESCC/premalignant lesions (Table 2). Older people had a higher probability of ESCC/premalignant lesions compared with younger ones ($P < 0.001$). Men had a higher risk than women ($P = 0.010$). Officers and peasants had a higher risk than people with other occupations ($P = 0.012$). Other esophageal lesions were also investigated. Age, sex, education, smoking, smoking exposure, drinking, mining exposure, dust exposure, gas exposure, foreign residency, and tea consumption were all associated with esophageal precursor lesions, while age, sex, occupation, labor intensity, income, smoking, smoking time, smoking exposure, drinking, mining exposure, mining exposure time, dust exposure, gas exposure, hypertension, diabetes, disease history, fried food consumption, fresh fruit consumption, and tea consumption were all associated with reflux esophagitis (Table 3).

A nomogram for the prediction of ESCC and premalignant lesions

We used the independent ESCC risk factors (age, sex, and occupation) to develop a predictive nomogram for the general population eligible for endoscopy screening. Three additional factors (labor intensity, income, mining exposure) were also included, considering their corresponding OR value in univariate analysis. Half of the cases were sampled by the Monte Carlo method, and a predictive nomogram was generated, which is shown in Figure 3A. For each patient, points were assigned for each of these demographic and medical factors (age, sex, education, occupation, labor intensity, income, and mining exposure), then a total score and a corresponding prediction of the probability of ESCC/premalignant lesions were calculated from the nomogram. An ROC curve was drawn to estimate the predictive accuracy of the nomogram, and the AUC (95% CI) was 0.749 (0.711–0.788; Figure 3B). The other half of the cases were used for validation, and the AUC (95% CI) was 0.742 (0.704–0.780) for the validation cohort, indicating a very good concordance (Figure 3C). A calibration curve gener-

Table 2 Univariate and multivariate analyses of risk factors for ESCC and premalignant lesions

	Univariate ^a			Multivariate		
	OR	95% CI	P	OR	95% CI	P
Age (years)			<0.001			<0.001
41–50 vs ≤40	1.06	(0.38, 4.44)		0.93	(0.33, 3.94)	
51–60 vs ≤40	3.76	(1.38, 15.48)		2.21	(0.78, 9.26)	
>60 vs ≤40	10.87	(3.61, 46.90)		3.88	(1.09, 18.50)	
Sex			0.009			0.010
Female vs Male	0.54	(0.32, 0.86)		0.62	(0.50, 0.79)	
Education			<0.001			
High school vs College	1.15	(0.59, 2.46)				
Middle school vs College	0.91	(0.47, 1.94)				
Primary school vs College	3.37	(1.58, 7.60)				
Occupation			<0.001			0.012
Peasant vs Worker	3.15	(0.76, 8.77)		1.98	(0.43, 6.44)	
Officer vs Worker	2.34	(1.37, 3.86)		2.86	(1.38, 5.75)	
Service vs Worker	0.44	(0.02, 2.01)		0.67	(0.04, 3.32)	
Retired vs Worker	4.76	(2.92, 7.61)		2.10	(1.02, 4.26)	
Others vs Worker	1.09	(0.06, 5.07)		0.73	(0.04, 4.01)	
Labor intensity			0.006			
Moderate vs Light	0.59	(0.37, 0.91)				
Heavy vs Light	0.39	(0.16, 0.80)				
Income (¥)			0.008			
2,001–3,000 vs ≤2,000	0.52	(0.32, 0.84)				
3,001–4,000 vs ≤2,000	0.44	(0.24, 0.76)				
>4,000 vs ≤2,000	0.47	(0.25, 0.84)				
Mining exposure			0.033			
Yes vs No	1.53	(1.04, 2.29)				
Hypertension			0.033			
No vs Yes	0.63	(0.42, 0.96)				
Tea consumption			0.038			
No vs Yes	0.60	(0.39, 0.97)				

Note: ^aPartial results of univariate analysis of ESCC and premalignant lesions risk factors; the complete results are found in Table S1.

Abbreviation: ESCC, esophageal squamous cell carcinoma.

ated by 1,000 repetitions of bootstrap sample corrections is also illustrated in Figure 3D.

Selecting eligible people for endoscopy screening

To shrink the population eligible for endoscopy screening in the high ESCC risk area, we assigned a risk score for ESCC and premalignant lesions to each participant in this study according to the nomogram. A 40% percentile cutoff was applied, and 3,126 out of 7,830 people were identified as at high risk. All ESCC patients (4/4) and three-quarters of patients with premalignant lesions (70/98) were successfully selected in this patient selection strategy (Table 4). Among the remaining 3,052 patients identified as the high-risk group, 784 also had other esophageal diseases (precursor lesions and reflux esophagitis). In the 4,704 people in the low-risk subgroup, 3,962 participants were completely healthy (84.2%), and no ESCC patients were misdiagnosed.

Discussion

Many independent cohort studies have indicated that screening of early ESCC could reduce its mortality, and the detection of precancerous lesions likewise has led to a reduction in the incidence of ESCC.^{12,15} However, there are no global or nationwide recommendations for ESCC screening.

Recently, most ESCC screening programs were conducted in the target population defined by several risk factors such as age and family history.^{12,14} Inefficient preselection of eligible people to include resulted in low cost-efficiency in ESCC screening programs.^{16,17} Endoscopic screening with a low true positive rate also caused poor compliance in the target population, according to our experience. A predictive model that provided improved individual risk assessment would facilitate people's decision-making on whether to accept an endoscopy examination.

In this study, age was associated with ESCC/premalignant lesions, precursor lesions, and reflux esophagitis. It is reason-

Table 3 Univariate analysis of risk factors for benign esophageal lesions

	Precursor lesions			Reflux esophagitis		
	OR	95% CI	P	OR	95% CI	P
Age (years)			<0.001			0.002
41–50 vs ≤40	1.59	(0.98, 2.78)		1.24	(0.94, 1.66)	
51–60 vs ≤40	2.42	(1.48, 4.23)		1.52	(1.14, 2.05)	
>60 vs ≤40	3.36	(1.80, 6.46)		1.56	(1.03, 2.34)	
Sex			<0.001			<0.001
Female vs Male	0.62	(0.50, 0.79)		0.23	(0.19, 0.28)	
Education			0.002			0.489
High school vs College	1.69	(1.18, 2.49)		0.94	(0.77, 1.17)	
Middle school vs College	1.37	(0.97, 2.02)		1.03	(0.84, 1.27)	
Primary school vs College	2.17	(1.38, 3.47)		0.89	(0.65, 1.20)	
Occupation			0.319			0.016
Peasant vs Worker	1.13	(0.48, 2.29)		0.71	(0.38, 1.23)	
Officer vs Worker	0.96	(0.71, 1.28)		1.07	(0.90, 1.28)	
Service vs Worker	1.21	(0.72, 1.92)		0.83	(0.57, 1.18)	
Retired vs Worker	1.39	(1.01, 1.86)		0.83	(0.65, 1.04)	
Others vs Worker	1.51	(0.71, 2.87)		0.38	(0.16, 0.75)	
Labor intensity			0.827			0.017
Moderate vs Light	0.97	(0.79, 1.19)		1.19	(1.04, 1.36)	
Heavy vs Light	0.92	(0.68, 1.22)		1.21	(1.01, 1.46)	
Mining exposure			<0.001			<0.001
Yes vs No	1.49	(1.23, 1.80)		1.95	(1.72, 2.21)	
Mining exposure time (hours/week)			<0.001			<0.001
1–40 vs 0	1.16	(0.82, 1.61)		1.94	(1.58, 2.37)	
41–60 vs 0	1.53	(1.23, 1.91)		1.96	(1.70, 2.28)	
>60vs 0	1.68	(1.26, 2.21)		1.97	(1.62, 2.40)	
Dust exposure			0.002			<0.001
Yes vs No	1.35	(1.12, 1.64)		1.83	(1.61, 2.08)	
Gas exposure			0.003			<0.001
Yes vs No	1.34	(1.11, 1.61)		1.70	(1.50, 1.92)	
Smoking			0.002			<0.001
Occasionally vs Yes	0.82	(0.55, 1.17)		0.97	(0.78, 1.20)	
Refrained vs Yes	1.19	(0.83, 1.66)		0.91	(0.72, 1.15)	
Never vs Yes	0.71	(0.58, 0.87)		0.41	(0.36, 0.48)	
Drinking			<0.001			<0.001
Occasionally vs Yes	0.69	(0.55, 0.87)		0.64	(0.55, 0.73)	
Refrained vs Yes	0.67	(0.53, 0.85)		0.35	(0.30, 0.42)	
Never vs Yes	1.38	(0.85, 2.12)		0.68	(0.47, 0.96)	
Smoke exposure			0.016			<0.001
Yes vs No	1.38	(1.06, 1.84)		1.69	(1.41, 2.04)	
Disease history			0.816			0.010
No vs Yes	0.98	(0.81, 1.18)		0.85	(0.75, 0.96)	
Hypertension			0.502			<0.001
No vs Yes	0.93	(0.75, 1.16)		0.76	(0.66, 0.87)	
Diabetes			0.056			0.051
No vs Yes	0.43	(0.21, 1.02)		0.73	(0.53, 1.02)	
Tea consumption			<0.001			<0.001
No vs Yes	1.80	(1.33, 2.50)		0.75	(0.64, 0.87)	
Fried food consumption			0.103			<0.001
Rarely or Never vs Yes	0.84	(0.68, 1.04)		0.79	(0.69, 0.91)	
Fresh fruit consumption			0.143			<0.001
Rarely or Never vs Yes	1.15	(0.95, 1.39)		1.45	(1.28, 1.65)	
Foreign residency			0.017			0.483
No vs Yes	0.60	(0.41, 0.91)		0.89	(0.66, 1.23)	

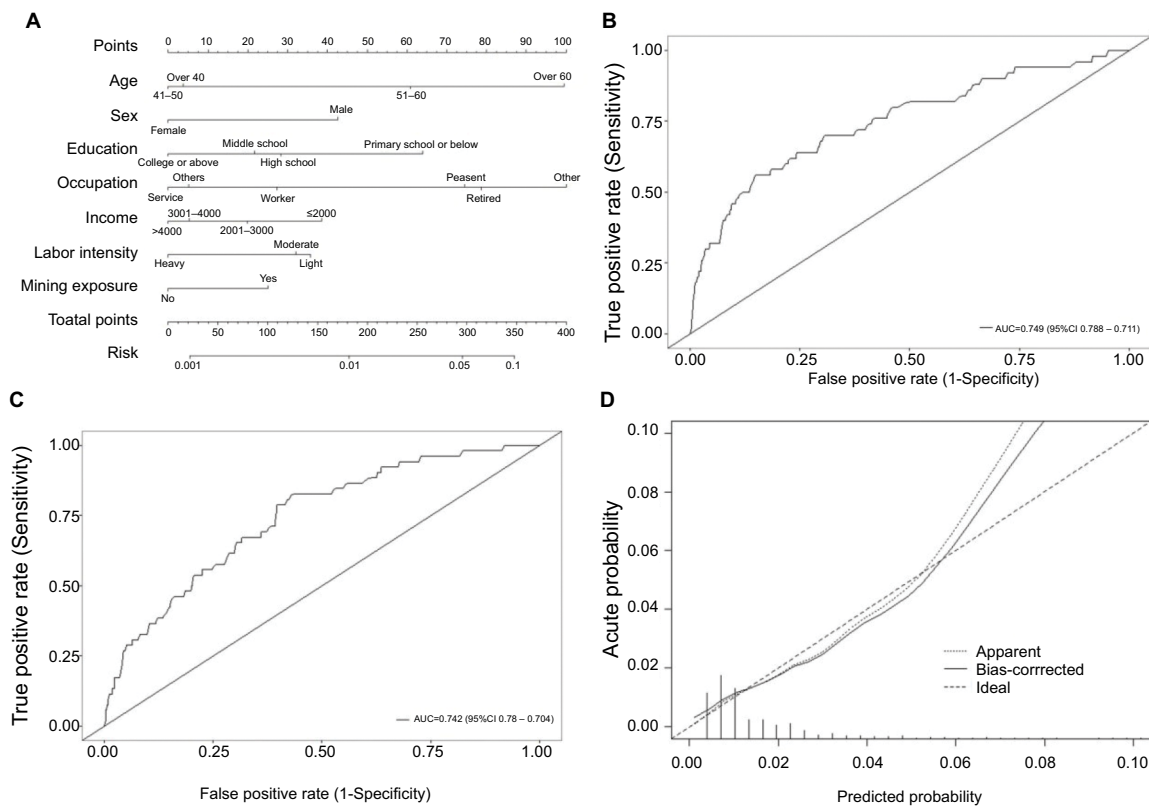


Figure 3 A nomogram for the prediction of ESCC and premalignant lesions.

Notes: Predictive nomogram of ESCC and premalignant lesions (**A**). ROC curve of this nomogram model (**B**). ROC curve of the validation of this nomogram model (**C**). Calibration curve of this nomogram model (**D**).

Abbreviations: ESCC, esophageal squamous cell carcinoma; ROC, receiver-operating characteristic.

Table 4 Sensitivity, specificity, and accuracy of patient selection strategy before endoscopy screening

Lesions	Sensitivity (%)	Specificity (%)	Positive prediction value (%)	Accuracy (%)
ESCC	100.0	53.7	0.1	53.7
Premalignant lesions	72.4	54.0	2.0	54.2
Precursor lesions	57.3	54.7	7.2	54.5
Reflux esophagitis	60.6	56.2	19.9	56.9

Abbreviation: ESCC, esophageal squamous cell carcinoma.

able that decreased immunity in older individuals resulted in a general increase in most diseases.¹⁸ Generally, people with lower income and who had received less education had a higher ESCC risk. Education and income were both associated with ESCC and premalignant lesions, which further supported the suggestion that ESCC is a poverty-related disease.¹⁹ Handan City has very considerable coal reserves, and many of our participants were involved in this industry. Hence, we included some occupational exposure factors in this study. Mining exposure was significantly associated with ESCC and premalignant lesions risk, and gas/dust exposures were risk factors for precancerous lesions/reflux esophagitis but not risk factors for ESCC/premalignant lesions. These results

suggested that employees of the mining industry were part of the high-risk population for ESCC, which was in accordance with previous reports.^{20,21} Thus, we advocated that additional medical attention should be addressed in this population.

Nomograms as risk estimators have shown promising potential in clinical trial design and interpretation and have been widely adopted in prognostic models.²²⁻²⁴ However, their application in primary health screening, such as cancer screening programs, is rare.^{25,26} In this study, we established a nomogram-based method to select the eligible population for endoscopic screening based on different risk factors. Through such a potential patient selection strategy, we could exclude 60% of people before endoscopic screening without missing

even one ESCC case. To our knowledge, this is the first study providing a nomogram to predict individual ESCC risk. These results suggest that the nomogram is a very powerful tool in individual risk assessment, one that could be used in ESCC screening and, most likely, in other primary health screening programs, as well.

There are also some limitations in our nomogram-based model. First, all nomograms were based on a retrospective single-center dataset, which would weaken the confidence of our risk prediction model and shrink its application range. Second, we used only demographic and lifestyle information to construct the model and did not include any biochemistry or hematology tests. Nomograms based on a dataset with more dimensions may provide a more accurate prediction, at a much greater cost. However, the aim of this study was not to deliver a perfect predictive tool but to develop a proper strategy for selecting an eligible population for endoscopic screening that was better than currently available approaches. For ESCC, a disease mainly occurring in developing areas, our model would be much available and affordable.

In conclusion, we provided a ready-to-use preclinical tool with the potential to select eligible people with high ESCC risk for endoscopy screening. This tool is not ideal but is very useful, and in the future, a multicentered prospective study with a larger population would provide a more reliable patient selection strategy.

Data availability

Additional supporting information may be found in the online version of this article.

Acknowledgments

This work was fully supported by National Natural Science Foundation of People's Republic of China (grant nos. 81302160 and 81272447), the Beijing Natural Science Foundation Program, the Research Foundation of Beijing Friendship Hospital, Capital Medical University (grant no. yyqdk2017-15) and the Scientific Research Key Program of the Beijing Municipal Commission of Education (grant no. KZ201410025024).

We would like to deeply appreciate the core members of the North China Esophageal Cancer Screening (NCECS) Study Group, whose names were not able to be fully displayed in the author list according to GPP3 guidelines. Their dedicated endoscopy operations were the most important intellectual contributions to this study, which could never be ignored.

Core member of the NCECS Study Group: Xuan Liu,¹ Qiaozhi Zhou,¹ Yue Jiao,¹ Ying Meng,¹ Ye Zong,¹ Qinggang Wang,¹ Yinglin Niu,¹ Hongmei Li,¹ Li Yu,¹ Haiying Zhao,¹ Wenyan Li,¹ Xun Yang,¹ Li Zhang,¹ Fandong Meng,¹ Haiyi Hu,¹ Yanhua Zhou,¹ Shuxia Yu,² Baohong Xu,³ Jianyu Hao,⁴ Aimin Wang,⁵ Zhen He,¹ Chuntao Liu,¹ Jian Wei,¹ Yue Tian,¹ Pei Zhang,¹ Zhonglin Yu¹

¹Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, People's Republic of China; ²Department of Gastroenterology, Daxing Hospital, Beijing 102600, People's Republic of China; ³Department of Gastroenterology, Luhe Hospital, Beijing 101100, People's Republic of China; ⁴Department of Gastroenterology, Beijing Chaoyang Hospital, Beijing 100020, People's Republic of China; ⁵Department of Gastroenterology, The 252nd Hospital of PLA, Baoding 071000, People's Republic of China.

Author contributions

JX and LM carried out the main analysis. JX, LM, S. Zhu, and S. Zhang conceived and designed the study. HZ and JW coordinated the cooperation between the National Clinical Research Center for Digestive Disease and the local community. ZZ and HL helped to perform statistical analysis. XS and her group performed the pathological diagnosis. QG, SW, and YZ helped to collect data. AW helped to analyze data. LM, JX, ZZ, HL, and S. Zhang drafted the manuscript. WL and FL instructed those conducting the endoscopy examinations. Y. Wu, PL, MJ, and Y. Wang monitored the endoscopy examinations. All other coauthors and the NCECS Study Group performed endoscopy examinations. All authors contributed to data analysis, drafting and revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work. S. Zhang is the guarantor of the article.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65(2):87–108.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin.* 2015;65(1):5–29.
3. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016;66(2):115–132.
4. Yu S, Zhang W, Ni W, et al. Nomogram and recursive partitioning analysis to predict overall survival in patients with stage IIB-III thoracic esophageal squamous cell carcinoma after esophagectomy. *Oncotarget.* 2016;7(34):55211–55221.
5. Luo LN, He LJ, Gao XY, et al. Evaluation of preoperative staging for esophageal squamous cell carcinoma. *World J Gastroenterol.* 2016;22(29):6683–6689.

6. Kumar V, Abbas A, Fausto N, Aster J. *Pathologic Basis of Disease*. Robbins and Cotran, editors. Philadelphia: Saunders Elsevier; 2010.
7. Morimoto M, Nishiyama K, Nakamura S, et al. Significance of endoscopic screening and endoscopic resection for esophageal cancer in patients with hypopharyngeal cancer. *Jpn J Clin Oncol*. 2010;40(10):938–943.
8. Chaber-Ciopinska A, Kiprian D, Kawecki A, Kaminski MF. Surveillance of patients at high-risk of squamous cell esophageal cancer. *Best Pract Res Clin Gastroenterol*. 2016;30(6):893–900.
9. PDQ. Screening and Prevention Editorial B. Esophageal Cancer Screening (PDQ(R)): Health Professional Version. *PDQ Cancer Information Summaries*. Bethesda (MD): National Cancer Institute; 2016.
10. Hu WS, Liu ZJ, Zhang JB, et al. Risk patterns of subsequent primary cancers following esophagectomy in early-stage thoracic esophageal squamous cell cancer patients. *Tumori*. 2015;101(3):328–333.
11. Lucenteforte E, Garavello W, Bosetti C, et al. Diet diversity and the risk of squamous cell esophageal cancer. *Int J Cancer*. 2008;123(10):2397–2400.
12. Wei WQ, Chen ZF, He YT, et al. Long-term follow-up of a community assignment, one-time endoscopic screening study of esophageal cancer in China. *J Clin Oncol*. 2015;33(17):1951–1957.
13. Yokoyama A, Oda J, Iriguchi Y, et al. A health-risk appraisal model and endoscopic mass screening for esophageal cancer in Japanese men. *Dis Esophagus*. 2013;26(2):148–153.
14. Domper Arnal MJ, Ferrández Arenas Á, Lanás Arbeloa Á. Esophageal cancer: risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol*. 2015;21(26):7933–7943.
15. Wani S, Das A, Rastogi A, et al. Endoscopic ultrasonography in esophageal cancer leads to improved survival rates: results from a population-based study. *Cancer*. 2015;121(2):194–201.
16. Gupta N, Bansal A, Wani SB, Gaddam S, Rastogi A, Sharma P. Endoscopy for upper GI cancer screening in the general population: a cost-utility analysis. *Gastrointest Endosc*. 2011;74(3):610–624.
17. Yang J, Wei WQ, Niu J, et al. Estimating the costs of esophageal cancer screening, early diagnosis and treatment in three high risk areas in China. *Asian Pac J Cancer Prev*. 2011;12(5):1245–1250.
18. Yaqoob P. Ageing alters the impact of nutrition on immune function. *Proc Nutr Soc*. 2017;76(3):347–351.
19. Lagergren J, Andersson G, Talbäck M, et al. Marital status, education, and income in relation to the risk of esophageal and gastric cancer by histological type and site. *Cancer*. 2016;122(2):207–212.
20. Santibañez M, Vioque J, Alguacil J, et al. Occupational exposures and risk of oesophageal cancer by histological type: a case-control study in eastern Spain. *Occup Environ Med*. 2008;65(11):774–781.
21. Vizecaino AP, Parkin DM, Skinner ME. Risk factors associated with oesophageal cancer in Bulawayo, Zimbabwe. *Br J Cancer*. 1995;72(3):769–773.
22. Kim SK, Chai YJ, Park I, et al. Nomogram for predicting central node metastasis in papillary thyroid carcinoma. *J Surg Oncol*. 2017;115(3):266–272.
23. 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(21):35311–35325.
24. 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(40):64757–64765.
25. Bass JL, Bhatia A, Boas FE, Sansary J, Rauch D. Validation of a body mass index nomogram for children as an obesity screening tool in young children. *Clin Pediatr*. 2006;45(8):718–724.
26. 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(1):69–72.

Supplementary material

Table S1 Univariate analysis of esophageal lesions risk factors

	Precursor lesions			Premalignant/malignant lesions			Reflux esophagitis		
	OR	95%CI	P	OR	95%CI	P	OR	95%CI	P
Age (years)			<0.001			<0.001			0.002
41–50 vs ≤40	1.59	(0.98, 2.78)		1.06	(0.38, 4.44)		1.24	(0.94, 1.66)	
51–60 vs ≤40	2.42	(1.48, 4.23)		3.76	(1.38, 15.48)		1.52	(1.14, 2.05)	
>60 vs ≤40	3.36	(1.80, 6.46)		10.87	(3.61, 46.90)		1.56	(1.03, 2.34)	
Gender			<0.001			0.009			<0.001
Female vs Male	0.62	(0.50, 0.79)		0.54	(0.32, 0.86)		0.23	(0.19, 0.28)	
Race			0.614			0.174			0.834
Others vs Han	0.71	(0.11, 2.29)		3.21	(0.52, 10.58)		1.09	(0.47, 2.20)	
Married			0.273			0.986			0.344
Others vs Married	1.59	(0.66, 3.25)		1.02	(0.06, 4.67)		0.72	(0.34, 1.38)	
Education			0.002			<0.001			0.489
High school vs College	1.69	(1.18, 2.49)		1.15	(0.59, 2.46)		0.94	(0.77, 1.17)	
Middle school vs College	1.37	(0.97, 2.02)		0.91	(0.47, 1.94)		1.03	(0.84, 1.27)	
Primary school vs College	2.17	(1.38, 3.47)		3.37	(1.58, 7.60)		0.89	(0.65, 1.20)	
Occupation			0.319			<0.001			0.016
Peasant vs Worker	1.13	(0.48, 2.29)		3.15	(0.76, 8.77)		0.71	(0.38, 1.23)	
Officer vs Worker	0.96	(0.71, 1.28)		2.34	(1.37, 3.86)		1.07	(0.90, 1.28)	
Service vs Worker	1.21	(0.72, 1.92)		0.44	(0.02, 2.01)		0.83	(0.57, 1.18)	
Retired vs Worker	1.39	(1.01, 1.86)		4.76	(2.92, 7.61)		0.83	(0.65, 1.04)	
Others vs Worker	1.51	(0.71, 2.87)		1.09	(0.06, 5.07)		0.38	(0.16, 0.75)	
Labor intensity			0.827			0.006			0.017
Moderate vs Light	0.97	(0.79, 1.19)		0.59	(0.37, 0.91)		1.19	(1.04, 1.36)	
Heavy vs Light	0.92	(0.68, 1.22)		0.39	(0.16, 0.80)		1.21	(1.01, 1.46)	
Income(¥)			0.119			0.008			0.232
2,001–3,000 vs ≤2,000	0.79	(0.62, 1.01)		0.52	(0.32, 0.84)		0.98	(0.83, 1.16)	
3,001–4,000 vs ≤2,000	0.75	(0.57, 0.98)		0.44	(0.24, 0.76)		1.15	(0.96, 1.38)	
>4,000 vs ≤2,000	0.75	(0.56, 0.99)		0.47	(0.25, 0.84)		1.01	(0.83, 1.22)	
Mining exposure			<0.001			0.033			<0.001
Yes vs No	1.49	(1.23, 1.80)		1.53	(1.04, 2.29)		1.95	(1.72, 2.21)	
Mining exposure time (hours/week)			<0.001			0.122			<0.001
1–40 vs 0	1.16	(0.82, 1.61)		1.39	(0.68, 2.62)		1.94	(1.58, 2.37)	
41–60 vs 0	1.53	(1.23, 1.91)		1.65	(1.04, 2.61)		1.96	(1.70, 2.28)	
>60 vs 0	1.68	(1.26, 2.21)		1.67	(0.89, 2.95)		1.97	(1.62, 2.40)	
Dust exposure			0.002			0.299			<0.001
Yes vs No	1.35	(1.12, 1.64)		1.23	(0.83, 1.84)		1.83	(1.61, 2.08)	
Gas exposure			0.003			0.610			<0.001
Yes vs No	1.34	(1.11, 1.61)		0.90	(0.60, 1.33)		1.70	(1.50, 1.92)	
Smoking			0.002			0.150			<0.001
Occasionally vs Yes	0.82	(0.55, 1.17)		1.00	(0.43, 2.03)		0.97	(0.78, 1.20)	
Refrained vs Yes	1.19	(0.83, 1.66)		1.90	(0.97, 3.48)		0.91	(0.72, 1.15)	
Never vs Yes	0.71	(0.58, 0.87)		0.86	(0.55, 1.33)		0.41	(0.36, 0.48)	
Drinking			<0.001			0.451			<0.001
Occasionally vs Yes	0.69	(0.55, 0.87)		1.08	(0.66, 1.74)		0.64	(0.55, 0.73)	
Refrained vs Yes	0.67	(0.53, 0.85)		0.93	(0.56, 1.54)		0.35	(0.30, 0.42)	
Never vs Yes	1.38	(0.85, 2.12)		2.06	(0.77, 4.62)		0.68	(0.47, 0.96)	
Smoking exposure			0.016			0.333			<0.001
Yes vs No	1.38	(1.06, 1.84)		0.79	(0.50, 1.30)		1.69	(1.41, 2.04)	
Smoking time >15 minutes			0.181			0.924			<0.001
Yes vs No	0.87	(0.70, 1.07)		0.98	(0.63, 1.48)		0.70	(0.61, 0.81)	

(Continued)

Table S1 (Continued)

	Precursor lesions			Premalignant/malignant lesions			Reflux esophagitis		
	OR	95%CI	P	OR	95%CI	P	OR	95%CI	P
Disease history			0.816			0.188			0.010
No vs Yes	0.98	(0.81, 1.18)		0.77	(0.52, 1.14)		0.85	(0.75, 0.96)	
Heart disease			0.207			0.807			0.840
No vs Yes	0.82	(0.61, 1.12)		0.92	(0.50, 1.89)		1.02	(0.83, 1.27)	
Diabetes			0.264			0.056			0.061
No vs Yes	1.41	(0.79, 2.87)		0.43	(0.21, 1.02)		0.73	(0.53, 1.02)	
Hypertension			0.502			0.033			<0.001
No vs Yes	0.93	(0.75, 1.16)		0.63	(0.42, 0.96)		0.76	(0.66, 0.87)	
Digestive disease			0.816			0.292			0.225
No vs Yes	1.03	(0.79, 1.36)		0.76	(0.47, 1.29)		1.12	(0.94, 1.34)	
Respiratory disease			0.536			0.342			0.079
No vs Yes	0.90	(0.66, 1.26)		0.74	(0.42, 1.42)		0.83	(0.68, 1.02)	
Tumor history			0.827			0.879			0.821
No vs Yes	0.98	(0.80, 1.20)		0.97	(0.64, 1.49)		1.02	(0.89, 1.16)	
Other disease history			0.507			0.831			0.664
No vs Yes	0.91	(0.71, 1.20)				0.038	0.96	(0.81, 1.15)	<0.001
Tea consumption			<0.001			0.038			<0.001
No vs Yes	1.80	(1.33, 2.50)		0.60	(0.39, 0.97)		0.75	(0.64, 0.87)	
Eating postures			0.644			0.069			0.463
Sit vs Squat	0.91	(0.47, 2.04)		0.80	(0.25, 4.87)		1.34	(0.80, 2.40)	
Stand vs Squat	1.93	(0.27, 8.65)		0.73	(0.88, 68.48)		0.87	(0.13, 3.47)	
High temperature food intake			0.095			0.561			0.511
No vs Yes	1.18	(0.97, 1.43)		1.13	(0.76, 1.69)		0.96	(0.85, 1.09)	
Fast eating			0.238			0.631			0.087
No vs Yes	1.12	(0.93, 1.35)		0.91	(0.61, 1.34)		0.90	(0.79, 1.02)	
Salted food consumption			0.449			0.654			0.544
Rarely or Never vs Yes	0.93	(0.77, 1.13)		1.10	(0.73, 1.67)		0.96	(0.85, 1.09)	
Fried food consumption			0.103			0.270			<0.001
Rarely or Never vs Yes	0.84	(0.68, 1.04)		0.78	(0.52, 1.22)		0.79	(0.69, 0.91)	
Chilli food consumption			0.295			0.187			0.724
Rarely or Never vs Yes	1.12	(0.91, 1.40)		1.36	(0.86, 2.24)		0.98	(0.85, 1.12)	
Moldy food consumption			0.751			0.83			0.916
No vs Yes	1.20	(0.44, 4.94)		0.80	(0.17, 14.18)		1.04	(0.54, 2.26)	
Leek food consumption			0.339			0.535			0.702
Rarely or Never vs Yes	0.88	(0.66, 1.15)				0.701	0.97	(0.81, 1.15)	
Bean food consumption			0.177			0.701			0.334
Rarely or Never vs Yes	0.83	(0.63, 1.08)		0.90	(0.50, 1.51)		0.92	(0.77, 1.09)	
Meat/egg/milk consumption			0.359			0.223			0.115
Rarely or Never vs Yes	0.83	(0.54, 1.22)		0.56	(0.17, 1.35)		0.81	(0.62, 1.05)	
Dry food consumption			0.133			0.198			0.404
Rarely or Never vs Yes	1.17	(0.95, 1.43)		1.32	(0.87, 2.05)		1.06	(0.93, 1.20)	
Maize consumption			0.177			0.075			0.530
Rarely or Never vs Yes	1.14	(0.94, 1.38)		1.43	(0.96, 2.12)		1.04	(0.92, 1.18)	
Fresh vegetable consumption			0.036			0.316			0.213
Rarely or Never vs Yes	10 ⁻⁶	(10 ⁻³⁸ , 0.22)		10 ⁻⁶	(10 ⁻⁵⁹ , 526.04)		0.50	(0.12, 1.41)	
Fresh fruit consumption			0.143			0.090			<0.001
Rarely or Never vs Yes	1.15	(0.95, 1.39)		1.40	(0.95, 2.10)		1.45	(1.28, 1.65)	
Sauerkraut consumption			0.288			0.667			0.656
No vs Yes	1.16	(0.88, 1.56)		0.89	(0.53, 1.57)		0.96	(0.81, 1.45)	
Water source			0.839			0.481			0.316
Natural vs Others	1.04	(0.67, 1.55)		1.34	(0.56, 2.70)		1.15	(0.87, 1.49)	
Foreign residency			0.017			0.137			0.483
No vs Yes	0.60	(0.41, 0.91)		0.53	(0.26, 1.26)		0.89	(0.66, 1.23)	

Cancer Management and Research

Dovepress

Publish your work in this journal

Cancer Management and Research is an international, peer-reviewed open access journal focusing on cancer research and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient. The manuscript management system is completely online and includes

a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/cancer-management-and-research-journal>