




Performance evaluation of four prediction models for risk stratification in gastric cancer screening among a high-risk population in China

Yue Hu^{1,2} · Haibiao Bao^{1,2} · Haifeng Jin^{1,2} · Jin Zhao^{1,2} · Yi Xu^{1,2} · Xuan Huang^{1,2} · Shan Liu³ · Bin Lu^{1,2} 

Received: 10 March 2021 / Accepted: 13 June 2021 / Published online: 21 June 2021
© The International Gastric Cancer Association and The Japanese Gastric Cancer Association 2021

Abstract

Objective Early detection of gastric cancer (GC) is a critical step for decreasing mortality. The aim of this study was to evaluate the performance of four prediction models for risk stratification in the screening of GC and precancerous lesions among a large, high-risk population in China.

Design This study was a retrospective analysis of data from the Provincial Gastric Cancer Screening Program (Zhejiang, China) spanning the period between October 2016 and April 2019, in which 97,541 individuals from the urban areas of 10 cities in Zhejiang province, China participated in this program. Demographic and clinical characteristics data were collected, and serum pepsinogens I and II, gastrin-17, and anti-*H. pylori* IgG antibody were detected. Participants were asked to voluntarily undergo gastroscopy. The performance of the ABC method, new ABC method, Tu's prediction model, and Li's prediction model, which stratified participants into low-, medium- and high-risk subgroups, were evaluated using the area under the receiver-operating characteristic (ROC) curve (AUC) and Youden index.

Results Among the participants, 6005 (3447 males and 2558 females, mean age of 58.35 years), voluntarily underwent gastroscopy. Overall, 72 (1.20%) GC cases (30 early and 42 advanced) and 2006 cases with precancerous lesions (270 atrophic gastritis, 1634 intestinal metaplasia, and 102 dysplasia/intraepithelial neoplasia) were identified. Notably, Li's prediction model achieved the greatest AUC and Youden index values (0.708 and 0.319, respectively) for predicting GC, and exhibited the greatest ability to detect precancerous lesions, especially intestinal metaplasia.

Conclusion Li's prediction model performs the best for risk stratification in the screening, detection, and diagnosis of GC and precancerous lesions, whereas the overall performance of the other three models is similar (www.chictr.org.cn, ChiCTR2100043363).

Keywords GC · *Helicobacter pylori* · Pepsinogen · Gastrin-17 · Prediction model

Abbreviations

GC Gastric cancer
ROC Receiver-operating characteristic
PG Pepsinogens
CFU Colony-forming units

G-17 Gastrin-17
UGI Upper gastrointestinal
PPIs Proton pump inhibitors
ELISA Enzyme-linked immunosorbent assay
AG Atrophic gastritis
IM Intestinal metaplasia
IEN Dysplasia/intraepithelial neoplasia
WHO World Health Organization
ESD Endoscopic submucosal dissection

✉ Bin Lu
drlvbin@163.com

¹ Department of Gastroenterology, First Affiliated Hospital of Zhejiang Chinese Medical University, 54 Youdian Road, Hangzhou 310006, China

² Key Laboratory of Digestive Pathophysiology of Zhejiang Province, First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China

³ Center of Clinical Evaluation, The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China

Introduction

Gastric cancer (GC) is among the most common malignancies globally, with particularly high incidence in China, Korea, and Japan [1, 2]. In China, GC has caused

approximately 498,000 deaths nationwide in the year 2015, accounting for nearly half of GC-related deaths on the globe [3, 4]. It has been thought that late detection and delay in diagnosis and treatment are strongly associated with the high mortality of GC, while early detection is critical to reduce the mortality and make a substantial improvement in the survival of GC patients [5–7]. Unfortunately, a majority of GC cases in China are diagnosed at its advanced stage and this has posed tremendous challenges in effective treatment of GC, led to poor prognosis as well as high mortality of GC nationwide [5, 7]. Currently, there is no established screening program for GC in China, and early GC (EGC) is mainly detected in opportunistic screening [4, 5, 8, 9]. Given the high incidence and mortality of GC in China, it is urgently needed to find an effective and feasible screening program for GC, especially to the population in the high-risk areas, where the incidence of GC is greater than 30/100,000.

Gastroscopy, along with a histological examination of gastric mucosal biopsies, is considered the gold standard for screening, detection, and early diagnosis of GC. In a massive screening program for GC, initial risk stratification is usually conducted to identify the individuals at high risk prior to gastroscopy. In fact, the past decade has witnessed substantial progress in the development of models to predict risk levels in relation to GC, after which individuals at higher risk were identified to undergo gastroscopy for detection and early diagnosis of GC. Of these models, the ABC method was originally developed by Miki et al., using assays of serum levels of pepsinogens (PG), including PGI and PGII, in combination with anti-*Helicobacter pylori* IgG antibody for initial risk assessment for GC, and has been shown beneficial to screening for GC in the Japanese population [10]. A number of previous studies have tested the ABC method in screening for GC in the Chinese population, and identified the limitation in defining atrophic PG status [9, 11]. Thus, a new ABC method was recommended [12]. Then, Tu and colleagues assessed the values of five serum markers, including PGI, PGII, the ratio of PGI/PGII (PGR), anti-*H. pylori* IgG antibody, and gastrin-17 (G-17), for stratifying the individual's risk of developing GC and identifying individuals at higher risk needed for further gastroscopy [9]. In comparison with the ABC method, Tu's prediction model showed better performance in predicting GC risk and identifying individuals who were at higher risk of developing GC and needed for diagnostic gastroscopy [9, 10]. It was of note that the participating individuals presented upper gastrointestinal (UGI) symptoms or had a family history of GC [9], and the performance of Tu's prediction model in the population without UGI symptoms remains to be investigated. Most recently, a new model for predicting the risk of GC in the screening for GC has developed in the Chinese asymptomatic population in a cross-sectional study [8]. Compared with the three previously published models (i.e. the ABC method, new ABC method, and Tu's prediction model), this prediction

model showed better performance and ability in identifying individuals with GC, and with its potential in detecting GC at early stages [8]. Although the strength of Li's prediction model has been demonstrated, the model has not been externally validated in a large population. To date, a direct efficacy comparison of these proposed models in a large-size population in a Chinese high-incidence area is lacking. In addition, gastric precancerous lesions, if left untreated, can progress to GC, and thus, prediction models for precancerous lesions are also needed to effectively identify the high-risk patients.

The incidence of GC was reportedly 34.71/100,000 during the period between 2010 and 2014 in registration areas of Zhejiang Province, China [13], higher than 30/100,000 for defining the high-risk areas of GC in China. In an effort to reduce the high mortality and to improve the 5-year survival rate for GC, the screening for GC and early diagnosis as a provincial program was initiated and conducted in Zhejiang Province in 2016. The aim of this present study, as a part of this program, was to evaluate the performance of the various prediction models described above for GC screening in a large, high-risk population in China.

Study subjects and methods

Study population

In this retrospective study, data were obtained from the Provincial Gastric Cancer Screening Program (Zhejiang, China) spanning from October 2016 to April 2019. A total of 97,541 individuals in the urban areas of 10 cities, Zhejiang province, China, were willing to participate in this program. All participants had no gastrointestinal symptoms or occasionally mild symptoms. The demographic and clinicopathological characteristics of all participants are listed in Table 1. The detailed information was presented in the supplementary material.

This retrospective study was reviewed and approved by the Ethics Committee of the First Affiliation Hospital of Zhejiang Chinese Medical University (2020-KL-020-01), and registered in www.chictr.org.cn (Chinese Clinical Trial Registry, ChiCTR2100043363). All study procedures involving humans were conducted in compliance with the 1964 Helsinki Declaration. Participants had provided their signed informed consent.

Data collection, serological tests, upper endoscopy and histopathology, and the four prediction models

The details of data collection, serological tests, upper endoscopy and histopathology, and the four prediction models were described in Supplementary Material.

Table 1 Demographic and serological characteristics of the study population, and clinicopathological features of individuals undergoing further gastroscopic examination

Characteristics	Individuals with gastroscopy (<i>n</i> = 6005)	Individuals without gastroscopy (91,536)	Total individuals (<i>n</i> = 97,541)
Sex			
Female	3447 (57.40%)	54,096 (59.10%)	57,543 (58.99%)
Male	2558 (42.60%)	37,440 (40.90%)	39,998 (41.01%)
Age (years)			
< 40	151 (2.51%)	3843 (4.20%)	3994 (4.09%)
40–49	797 (13.27%)	14,496 (15.84%)	15,293 (15.68%)
50–59	2286 (38.07%)	33,484 (36.58%)	35,770 (36.67%)
60–69	2064 (34.37%)	28,573 (31.22%)	30,637 (31.41%)
70–79	657 (10.94%)	9747 (10.65%)	10,404 (10.67%)
≥ 80	50 (0.83%)	1393 (1.52%)	1443 (1.48%)
Mean (SD)	58.35 (4.24)	53.18 (10.43)	53.50 (10.23)
PGI (ng/mL)			
Median (range)	149.55 (79.48–165.00)	113.69 (84.78–155.76)	113.84 (84.53–156.27)
PGII (ng/mL)			
Median (range)	23.87 (8.57–27.04)	8.62 (5.64–14.32)	8.86 (5.72–15.1)
PGI/II ratio			
> 7	2138 (35.60%)	79,562 (86.92%)	81,700 (83.76%)
3–7	3393 (56.50%)	10,831 (11.83%)	14,224 (14.58%)
< 3	474 (7.89%)	1143 (1.25%)	1617 (1.66%)
G-17 (pmol/L)			
Median (range)	14.36 (0.97–17.97)	2.60(1.29–6.16)	2.68 (1.28–6.50)
Anti-<i>H. pylori</i> IgG antibody			
Positive	3568 (59.42%)	37,616 (41.09%)	41,184 (42.22%)
Negative	2437 (40.58%)	53,920 (58.91%)	56,357 (57.78%)
Risk stratification by Li's prediction model			
Low risk (Scores 0–11)	3394 (56.52%)	60,937 (66.57%)	64,331 (65.95%)
Medium risk (Scores 12–16)	2072 (34.50%)	26,245 (28.67%)	28,317 (29.03%)
High risk (Scores 17–25)	539 (8.98%)	4354 (4.76%)	4893 (5.02%)
Histological assessment			
Within normal limit/chronic inflammation	3924 (65.35%)	N/A	N/A
Precancerous lesions ^a	2006	N/A	N/A
AG	270	N/A	N/A
IM	1634	N/A	N/A
Dysplasia/IEN	102	N/A	N/A
GC ^b	72 (1.20%)	N/A	N/A
Malignancies other than GC ^c	11 (0.18%)	N/A	N/A

PGI or PGII pepsinogen I or II, G-17 gastrin-17, SD standard deviation, GC gastric cancer, N/A not applicable, AG atrophic gastritis, IM intestinal metaplasia, IEN intraepithelial neoplasia

^aMore than one lesion existed in some cases

^b72 (30 early, and 42 advanced) GC cases

^cFive cases with esophageal squamous cell carcinoma and six with gastroesophageal junction adenocarcinoma

Statistical analysis

Statistical analysis was carried out using SPSS software (IBM, Armonk, NY, USA). Demographical and clinical characteristics of the study population were described as

frequencies for categorical variables. Continuous variables were expressed as Mean ± standard deviation (SD) or median (interquartile range). A risk score was assigned to each participant after summing up all score values. Receiver-operating characteristic curves (ROCs) were plotted, and

the area under the ROC curve (AUC), with 95% confidence intervals (CI), was calculated. The Youden index of each model was calculated using the formula as follows: Youden index = (Sensitivity + Specificity) - 1. The values of the AUC and Youden index were used for the evaluation of the performance of the four models. A two-sided P value < 0.05 was considered statistically significant. All authors had access to the study data and reviewed and approved the final manuscript.

Results

Demographic, serological, and clinicopathological characteristics of the study population

A total of 97,541 individuals from the urban areas of the participating cities who fulfilled the criteria participated in this study and the four models were used for risk assessment of GC. Among these participants with initial serological analyses for GC screening, 6005 individuals voluntarily underwent further gastroscopic examination in the screening program, accounting for 6.16% of all participants. The demographic and serological characteristics of the study population and clinicopathological features of the gastric mucosal biopsies of individuals with gastroscopy are summarized in Table 1. Among those undergoing gastroscopy, 72 GC cases (1.20%) were detected, in which 30 (41.67%) were diagnosed as early GC, and 42 (58.33%) were diagnosed as advanced GC. Moreover, precancerous lesions, including, were detected in 2006 participants (270 AG, 1634 IM and 102 dysplasia/IEN) undergoing gastroscopy (Table 1). In addition, five cases of esophageal squamous cell carcinoma (ESCC) and six cases of gastroesophageal junction adenocarcinoma (GEJAC) were identified as malignancies other than true gastric adenocarcinoma (GAC).

The detected GC patients were treated with either endoscopic submucosal dissection (ESD) for early GC or surgery for advanced GC. For the individuals identified to have precancerous lesions and tested positive for *H. pylori* infection, they were treated with *H. pylori* eradication therapy, and then with [14].

Performance of various models for predicting GC

The four prediction models were assessed for their performance in the screening for GC. The numbers and proportions of cases with GC, including early and advanced GC, among subgroups of the various prediction models are summarized in Table 2. There were significant differences in the detection rates of GC among the subgroups of Li's prediction model ($P < 0.001$) and the new ABC method ($P = 0.002$), whereas the ABC method and Tu's

prediction model did not show a significant difference in the GC detection rates among the subgroups (Table 2). There were significant differences in the detection rates of predicting early GC among the subgroups with Li's prediction model (both $P < 0.001$) and Tu's prediction model ($P = 0.039$), but not with the ABC method and new ABC method (Table 2). The detection rates for advanced GC were significantly different among the subgroups with Li's prediction model ($P < 0.001$) and the new ABC method ($P = 0.001$), but not with the ABC method and Tu's prediction method (Table 2). Notably, the detection rates were significantly different among the subgroups for both early and advanced GC with Li's prediction model, which detected 75.0% (54/72) of GC cases: 76.7% (23/30) of early GC cases, and 73.8% (31/42) of advanced GC cases (Table 2).

To further assess the performance of the four models, the ROCs were plotted (Fig. 1), and their AUCs and the Youden indexes were calculated (Table 3). ROCs yielded an AUC of 0.708 (95% CI 0.643–0.773) for Li's prediction model ($P < 0.001$), which was higher than that of 0.566 [95% CI 0.496–0.636] for the ABC method ($P = 0.054$), 0.496 [95% CI 0.420–0.572] for the new ABC method ($P = 0.905$), or 0.534 [95% CI 0.463–0.605] for Tu's prediction model ($P = 0.318$) (Table 3, Fig. 1). Li's prediction model had significantly better discrimination ability to predict GC in comparison with the ABC method, the new ABC method, or Tu's prediction model (all $P < 0.001$). In addition, Li's prediction model had a Youden index of 0.319, which was higher than that of the original ABC method (0.103), the new ABC method (0.103), or Tu's prediction model (0.079) (Table 3). Thus, Li's prediction model had the highest AUC and the Youden index values among the evaluated prediction models.

Performance of various models for predicting precancerous lesions

Overall, precancerous lesions, including 270 AG, 1634 IM, and 102 dysplasia/IEN, were detected in 2006 of the 6005 participants undergoing gastroscopic examination; some of whom had more than one precancerous lesion. The performance of the four models in predicting precancerous lesions, including AG, IM, and dysplasia/IEN, was evaluated. The numbers of detected cases with precancerous lesions and their proportions classified by the ABC method, new ABC method, Tu's prediction model, and Li's prediction model are shown in Table 4. The detection rates of all these precancerous lesions in Li's prediction model and Tu's prediction model were significantly different among the three risk-stratified subgroups ($P < 0.001$ for AG, IM and dysplasia/IEN).

Table 2 Comparison of detected gastric cancer cases and their proportion by risk-stratified subgroups in the four evaluated models

Model	Total (N=6005)	GC (N=72, %)	Early GC (N=30, %)	Advanced GC (N=42, %)
ABC method (model 1)				
Group A (PG ^{nor} /HP ⁻)	1994	19 (0.95)	8 (0.40)	11 (0.55)
Group B (PG ^{nor} /HP ⁺)	3119	35 (1.12)	15 (0.48)	20 (0.64)
Group C (PG ^{low} /HP ⁺)	449	7 (1.56)	1 (0.22)	6 (1.34)
Group D (PG ^{low} /HP ⁻)	443	11 (2.48)	6 (1.35)	5 (1.13)
<i>P</i> value		0.050	0.055	0.206
New ABC method (model 2)				
Group A (PG ^{nor} /G-17 ^{nor})	2303	34 (1.48)	13 (0.56)	21 (0.91)
Group B (PG ^{nor} /G-17 ^{abn})	2810	20 (0.71)	10 (0.36)	10 (0.36)
Group C (PG ^{low} /G-17 ^{nor})	317	4 (1.26)	3 (0.95)	1 (0.32)
Group D (PG ^{low} /G-17 ^{abn})	575	14 (2.43)	4 (0.70)	10 (1.74)
<i>P</i> value		0.002	0.379	0.001
Tu's prediction model (model 3)				
Low risk (scores ≤2)	907	4 (0.44)	1 (0.11)	3 (0.33)
Medium risk (scores 3–13)	2070	32 (1.55)	16 (0.77)	15 (0.72)
High risk (scores ≥14)	3028	38 (1.25)	13 (0.43)	24 (0.79)
<i>P</i> value		0.090	0.039	0.072
Li's prediction model (model 4)				
Low risk (scores 0–11)	3394	18 (0.53)	7 (0.21)	11 (0.32)
Medium risk (scores 12–16)	2072	27 (1.30)	8 (0.39)	19 (0.92)
High risk (scores 17–25)	539	27 (5.01)	15 (2.78)	12 (2.23)
<i>P</i> value		<0.001	<0.001	<0.001

For models 1–2, the risk for GC was stratified into low risk (Group A), medium risk (Group B), and high risk (Groups C and D); For model 3, the risk was stratified into low-risk (≤2), medium-risk (3–13), and high-risk (≥14); For model 4, Age, gender, PGR, G-17, *H. pylori* infection, fried food, and pickled food were included, with each variable assigned to a score value. The risk for GC was stratified into low-risk (0–11), medium-risk (12–16), and high risk (17–25)

A *P* value denotes the level of statistical significance of comparison among all subgroups

PG^{nor} or PG^{low} normal or low serum levels of pepsinogen I (PGI) and ratio of PGI/PGII in models 1 and 2 or normal or low serum levels of PGI in model 3, G-17^{nor} G-17^{abn} normal or low serum levels of gastrin-17 in models 2 and 3, HP⁺ or HP⁻ positive or negative for anti-*H. pylori* IgG antibody

Discussion

This is the first study, to the best of our knowledge, in which the previously published four prediction models were assessed in parallel for their performance in screening GC among a large-size population in a Chinese high-incidence area. The screening program enrolled 97,541 individuals, of which 6005 individuals voluntarily participated in further gastroscopy. The major novel findings are summarized as follows: (1) GC was detected in 72 (1.20%) of the 6005 participants: 30 (41.67%) with early and 42 (58.33%) with advanced GC; (2) One or more precancerous lesions were identified in 2006 cases: 270 with AG, 1634 with IM, and 102 with dysplasia/IEN; (3) Li's prediction model was the most accurate model for risk stratification of GC, as evidenced mainly by its discriminatory detection rates for overall, early and advanced GC among subgroups, the highest values of AUC and Youden index among various models,

while the performance of the other three models was similar; and (4) Li's prediction model showed better performance in risk stratification of precancerous lesions (AG, IM, and dysplasia/IEN) in comparison with the other three models. These findings are of particular importance, as they confirm that Li's prediction model is an accurate and reliable method for the initial identification of individuals at higher risk who should undergo gastroscopy for further diagnostic examinations in high-incidence areas in China.

A screening program for GC has not been available in China nationwide, even in the high-incidence areas (i.e. incidence of GC > 30/100 000), where a majority of GC cases are initially detected or diagnosed at its late stage. The present study was conducted to evaluate in a high-risk Chinese population the efficacy of the four published prediction models derived from different populations on the basis of various serological markers and other risk factors and to identify a reliable model to be used for identifying

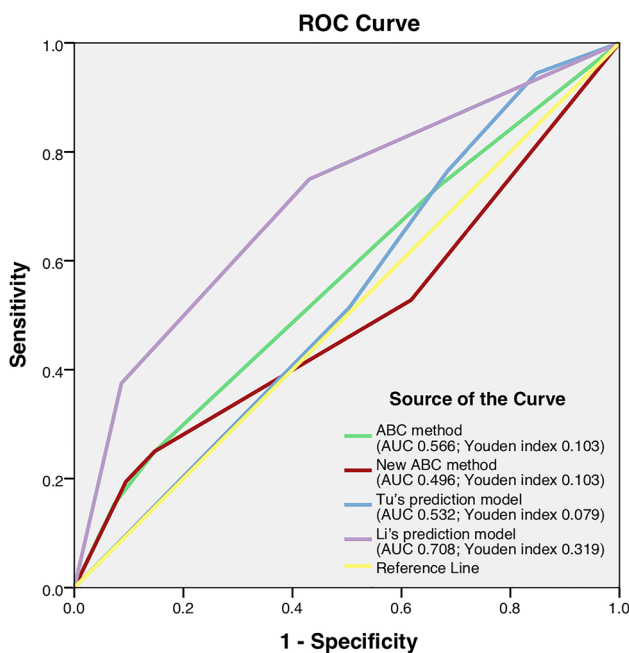


Fig. 1 Receiver-operating characteristic (ROC) curves and Youden indexes for the evaluation of discrimination ability of the four prediction models. ROC curves were generated for the four prediction models, including the original ABC method, new ABC method, modified ABC method, and Li's prediction model. The area under the ROC curve (AUC) and the Youden indexes were used for the performance evaluation of the four models. The AUC and the Youden index values are the highest for Li's prediction model among the four models

individuals at higher risk who need further diagnostic gastroscopy. This is particularly important, as both endoscopists and facilities are limited to sever the very large population in China, especially in the high incidence areas, where dietary habits and local foods are likely contributing factors in the transformation from precancerous lesions to gastric malignancies [15]. In the present study, the participants came from geographically diverse, high-incidence areas (GC incidence of 34.71/100,000). A total of 72 GC cases were detected and pathologically confirmed in this study. Of these cases, 30 patients diagnosed as having early GC subsequently received ESD procedures, which are expected to lead to good clinical outcomes. In addition to the 72 GC cases, gastric precancerous lesions were identified a larger number of cases, i.e. 2006 in total (270 AG, 1634 IM, and 102 dysplasia/IEN).

H. pylori infection is known to be a major cause of precancerous lesions. So, these patients with precancerous lesions were scheduled to take a further test for *H. pylori* infection. For patients confirmed positive for *H. pylori* infection, they were treated with an eradication regimen, and followed up regularly, according to the consensus and guideline in China (Shanghai, 2017; Changsha, 2014), to not only decrease the risk of GC development but also early detect GC.

Most recently, Li's team performed a multicenter cross-sectional study of the Chinese high-risk population, including 14,929 eligible individuals aged 40–80 years, and developed the prediction rule to be used for initial identification of higher-risk individuals for further gastroscopy [8]. According to Li's prediction model, the diagnostic gastroscopy was applied to the study population in medium-risk (scores 12–16) and high-risk groups (scores 17–25), 70.3% of early GC cases and 70.8% of total GC cases regardless of the stages were detected [8]. In the present study, Li's prediction model detected 75.0% of GC cases, 76.7% of early GC cases, and 73.8% advanced GC cases according to the recommendation of this model, which were higher than those obtained with the ABC method, the new ABC method, and Tu's prediction model. Our results were basically consistent with that reported by Li's team [8]. These findings obtained using Li's prediction model are encouraging and helpful for GC screening, and demonstrate that Li's prediction model is a reliable screening tool to be used for identifying individuals at higher risk who require further diagnostic gastroscopy. Moreover, based on the detection rates obtained with Li's prediction model in the individuals who voluntarily underwent gastroscopy, it would be estimated that the total GC case number among 97,541 individuals would be 954, including 341 (64,331 × 0.53%), 368 (28,317 × 1.3%), and 245 (4893 × 5.01%), respectively, in those with low-, medium- and high-risk (Tables 1 and 5). Thus, using Li's prediction model, approximately two-thirds of GC cases, (64.3% (613/954) in total, 64.5% (245/380) of early GC, and 64.1% (368/ 574) of advanced GC), would be detected among the 97,541 individuals in our initial screening program, which would avoid about two thirds (66.0% (64,331/97541) of endoscopic procedures (Table 5).

Our study has potential limitations. First, the population in diverse geographical areas of Zhejiang Province in Southeastern China was studied for evaluating the

Table 3 The area under the receiver-operating characteristic curve and Youden index for each of the four models

Model	AUC (95% CI)	Youden index	Sensitivity	Specificity	P value
ABC method	0.566 (0.496–0.636)	0.103	0.250	0.853	0.054
New ABC method	0.496 (0.420–0.572)	0.103	0.250	0.853	0.905
Tu's prediction model	0.528 (0.467–0.589)	0.097	0.944	0.152	0.343
Li's prediction model	0.708 (0.643–0.773)	0.319	0.750	0.569	<0.001

AUC area under the receiver operating characteristic curve, CI confidence interval

Table 4 Comparison of detected precancerous lesions and their proportions by subgroups applied to gastroscopies in the four evaluated models

Model	Total (N=6005)	AG (N=270, %)	IM (N=1634, %)	Dysplasia/IEN (N=102, %)
ABC method (model 1)				
Group A (PG ^{nor} /HP ⁻)	1994	61 (3.06)	421 (21.11)	31 (1.55)
Group B (PG ^{nor} /HP ⁺)	3119	163 (5.23)	943 (30.23)	51 (1.64)
Group C (PG ^{low} /HP ⁺)	449	26 (5.79)	171 (38.08)	8 (1.78)
Group D (PG ^{low} /HP ⁻)	443	20 (4.51)	99 (22.35)	12 (2.71)
<i>P</i> value		0.002	<0.001	0.348
New ABC method (model 2)				
Group A (PG ^{nor} /G-17 ^{nor})	2303	92 (3.99)	666 (28.92)	40 (1.74)
Group B (PG ^{nor} /G-17 ^{abn})	2810	132 (4.70)	698 (24.84)	42 (1.49)
Group C (PG ^{low} /G-17 ^{nor})	317	11 (3.47)	108 (34.07)	1 (0.32)
Group D (PG ^{low} /G-17 ^{abn})	575	35 (6.09)	162 (28.17)	19 (3.30)
<i>P</i> value		0.123	<0.001	0.004
Tu's prediction model (model 3)				
Low risk (scores ≤2)	907	33 (3.64)	181 (19.96)	14 (1.54)
Medium risk (scores 3–13)	2070	69 (3.33)	546 (25.07)	40 (1.93)
High risk (scores ≥14)	3028	168 (5.55)	934 (30.85)	48 (1.59)
<i>P</i> value		<0.001	<0.001	0.594
Li's prediction model (model 4)				
Low risk (scores 0–11)	3394	114 (3.36)	739 (21.77)	48 (1.41)
Medium risk (scores 12–16)	2072	121 (5.84)	659 (31.81)	35 (1.69)
High risk (scores 17–25)	539	35 (6.49)	236 (43.78)	19 (3.53)
<i>P</i> value		<0.001	<0.001	<0.001

For models 1–2, the risk for GC was stratified into low risk (Group A), medium risk (Group B), and high risk (Groups C&D); For model 3, the risk was stratified into low-risk (≤2), medium-risk (3–13), and high-risk (≥14); For model 4, Age, gender, PGR, G-17, *H. pylori* infection, fried food, and pickled food were included, with each variable assigned to a score value; The risk for GC was stratified into low-risk (0–11), medium-risk (12–16), and high risk (17–25)

A *P* value denotes the level of statistical significance of comparison among the subgroups

PG^{nor} or PG^{low} normal or low serum levels of pepsinogen I (PGI) and ratio of PGI/PGII in models 1 and 2, or normal or low serum levels of PGI in model 3, G-17^{nor} G-17^{abn} normal or low serum levels of gastrin-17 in models 2 and 3, HP⁺, or HP⁻ positive or negative for anti-*H. pylori* IgG antibody. AG atrophic gastritis, IM intestinal metaplasia, IEN intraepithelial neoplasia

performance of prediction models. Due to geographic variations in risk factors for GC, the performance and efficacy of the prediction models may vary in other regions in China. Second, a selection bias between with and without gastroscopy occurred, as individuals with gastroscopy were older and lower level of PG ratio than those without gastroscopy, which may influence the prevalence of GC. To reduce the bias in this real-world study, the results of GC detection rate were expressed as percentages and 95% CI, and the estimated number of cases in the corresponding CI had been calculated (Table 5). Third, some other prediction models have been reported but not evaluated for their performance in the present study. For instance, in a recent study by Zhan's team, serological tests of *H. pylori* IgG antibody and PGs, in combination with a family history of GC, were used for risk stratification and identification of individuals for further gastroscopy and gastric mucosal biopsies [16]. In this study, all individuals in

Groups C and D were mandatorily referred for a further gastroscopy, while a proportion of individuals in Groups A and B, consisting of individuals with a family history of GC in their first-degree relatives (FDR) and those selected at random from the remaining individuals, were referred for undergoing gastroscopy [16]. It was of note that the detection rate of GC in the study was 1.6% (GC cases per 100 gastroscopies) and 1.8% (GC cases per 1000 serological tests) [16]. In the present study, we yielded a detection rate of 1.20/100 gastroscopies (72/6005 gastroscopies). Actually, a number of previous studies have identified a history of GC in FDRs as a risk factor for both GC and gastric precancerous lesions [17–20]. Fourth, serology for anti-*H. pylori* IgG antibody was used for an evaluation of *H. pylori* status in the present study, as the test is less costly and more feasible than the ¹³C-urea breath test, which seems to be more accurate than serology. Fifth, it has to be pointed out that 25.0% of GC cases would be

Table 5 Estimated gastric cancer cases among the 97,541 individuals using Li's prediction model

	GC Detection rate (%)	Cases detected in individuals with gastroscopy <i>N</i> =6005	Cases estimated in individuals without gastroscopy <i>N</i> =91,536)	Cases estimated in total Individuals <i>N</i> =97,541
Low risk (scores 0–11)		<i>n</i> = 3394	<i>n</i> = 60,937	<i>n</i> = 64,331
GC	0.53 (0.31–0.84)	18	323 (189–512)	341 (199–540)
Early GC	0.21 (0.08–0.42)	7	128 (49–256)	135 (51–270)
Advanced GC	0.32 (0.16–0.58)	11	195 (98–353)	206 (103–373)
Medium risk (scores 12–16)		<i>n</i> = 2072	<i>n</i> = 26,245	<i>n</i> = 28,317
GC	1.30 (0.86–1.89)	27	341 (226–496)	368 (243–535)
Early GC	0.39 (0.17–0.76)	8	101 (44–197)	109 (48–212)
Advanced GC	0.92 (0.55–1.43)	19	240 (143–373)	259 (155–403)
High risk (scores 17–25)		<i>n</i> = 539	<i>n</i> = 4354	<i>n</i> = 4893
GC	5.01 (3.33–7.20)	27	218 (145–313)	245 (163–352)
Early GC	2.78 (1.57–4.55)	15	121 (68–198)	136 (77–223)
Advanced GC	2.23 (1.16–3.86)	12	97 (50–168)	109 (57–189)

Data in parentheses are 95% confidence intervals

missed in the study population (*n* = 6005) on the basis of the risk stratification using Li's prediction model in the present study. As such, other risk factors should be considered to improve the performance. For instance, genetic variations, such as *CYP1A1* and *cadherin one* genes, have been reported with a relationship to the development of GC or GC susceptibility [21], integrating genetic factors into the prediction model might further improve predictive ability.

In conclusion, Li's prediction model is an accurate and reliable model with the best performance among the four evaluated prediction models in stratifying the risk for GC and identifying the higher risk individuals for further gastroscopy in the screening, detection, and early diagnosis of GC. In addition, its discrimination power for GC and precancerous lesions is also greater than the other three models. Thus, Li's prediction model is a feasible and reliable tool in GC screening of the Chinese population. As screening gastroscopy for GC in the general population has not been covered by medical insurance in China nationwide, along with the limited endoscopists as well as facilities, Li's prediction model should be recommended for risk stratification to identify those individuals who are at higher risk of GC and need for further diagnostic gastroscopy and gastric mucosal biopsy.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10120-021-01204-6>.

Author contributions BL designed and supervised the study including all data collection and analysis; YH performed most of the investigation, including data collection and analysis, and wrote the manuscript. HB, HJ, YX, JZ and XH conduct the patients' gastroscopy and assisted

with the data collection. SL provided guidance in data analysis. All authors have read and approved the manuscript.

Funding This study was supported by the Natural Science Foundation of Zhejiang province (No. LQ21H030001), and Natural Science Foundation of China (No. 81970470).

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359–86.
2. Hahm MI, Choi KS, Lee HY, Jun JK, Oh D, Park EC. Who participates in the gastric cancer screening and on-time rescreening in the National Cancer Screening Program? A population-based study in Korea. *Cancer Sci*. 2011;102:2241–7.
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. 2016;66:115–32.
4. Nie Y, Wu K, Yu J, Liang Q, Cai X, Shang Y, et al. A global burden of gastric cancer: the major impact of China. *Expert Rev Gastroenterol Hepatol*. 2017;11:651–61.
5. Zong L, Abe M, Seto Y, Ji J. The challenge of screening for early gastric cancer in China. *Lancet*. 2016;388:2606.
6. Leja M, You W, Camargo MC, Saito H. Implementation of gastric cancer screening—the global experience. *Best Pract Res Clin Gastroenterol*. 2014;28:1093–106.
7. Orditura M, Galizia G, Sforza V, Gambardella V, Fabozzi A, Laterza MM, et al. Treatment of gastric cancer. *World J Gastroenterol*. 2014;20:1635–49.
8. Cai Q, Zhu C, Yuan Y, Feng Q, Feng Y, Hao 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.
9. Tu H, Sun L, Dong X, Gong Y, Xu Q, Jing J, et al. A serological biopsy using five stomach-specific circulating biomarkers for gastric cancer risk assessment: a multi-phase study. *Am J Gastroenterol*. 2017;112:704–15.
 10. Miki K. Gastric cancer screening by combined assay for serum anti-*Helicobacter pylori* IgG antibody and serum pepsinogen levels—“ABC method.” *Proc Jpn Acad Ser B*. 2011;87:405–14.
 11. Gao D, Zhao J, Yu X, Zhang B, Kong M, Fang Y, et al. Clinical values of gastrin 17 and pepsinogen in gastric cancer and precancerous lesion screening. *Biomed Res*. 2017;28:6140–3.
 12. Chinese Society of Digestive Endoscopy. Consensus on screening and endoscopic diagnosis and treatment of early gastric cancer in China (Changsha, 2014). *Zhong hua Xiao Hua Nei Jing Za Zhi*. 2014;31:361–77.
 13. Ding P, Du L, Li H, Zhang X, Zhang J, et al. Incidence and mortality of stomach cancer in Zhejiang cancer registration areas, 2010–2014. *China Cancer*. 2019;28:93–9.
 14. Fang JY, Du YQ, Liu WZ, Ren JL, Li YQ, Chen XY, et al. Chinese consensus on chronic gastritis (2017, Shanghai). *J Dig Dis*. 2018;19(4):182–203.
 15. You WC, Zhang L, Gail MH, Chang YS, Liu WD, Ma JL, et al. Gastric dysplasia and gastric cancer: *helicobacter pylori*, serum vitamin C, and other risk factors. *J Natl Cancer Inst*. 2000;92:1607–12.
 16. Ji L, Liu Z, Zhou B, Cai Y, An F, Wang L, et al. Community-based pilot study of a screening program for gastric cancer in a chinese population. *Cancer Prev Res*. 2020;13:1–10.
 17. Yaghoobi M, McNabb-Baltar J, Bijarchi R, Hunt RH. What is the quantitative risk of gastric cancer in the first-degree relatives of patients? A meta-analysis. *World J Gastroenterol*. 2017;23:2435–42.
 18. Bernini M, Barbi S, Roviello F, Scarpa A, Moore P, Pedrazzani C, et al. Family history of gastric cancer: a correlation between epidemiologic findings and clinical data. *Gastric Cancer*. 2006;9:9–13.
 19. Correa P. Human gastric carcinogenesis: a multistep and multifactorial process—First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer Res*. 1992;52:6735–40.
 20. Song M, Camargo MC, Weinstein SJ, Best AF, Männistö S, Albanes D, et al. Family history of cancer in first-degree relatives and risk of gastric cancer and its precursors in a Western population. *Gastric Cancer*. 2018;21:729–37.
 21. Han F, Wang X, Wang X, Luo Y, Li W. Meta-analysis of the association of CYP1A1 polymorphisms with gastric cancer susceptibility and interaction with tobacco smoking. *Mol Biol Rep*. 2012;39:8335–44.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.