Clin. Lab. 2026;72:XXX-XXX
©Copyright

ORIGINAL ARTICLE

A Predictive Model for Carcinogenic Risk in Chronic Atrophic Gastritis Using Serology and Gastroscopy

Li Zhang, Xin-Yu Yang

Department of Traditional Chinese Medicine, Dingxi People's Hospital, Dingxi City, Gansu Province, China

ABSTRACT

Background: **This study aimed to develop a simplified predictive model for assessing the risk of carcinogenic transformation in patients with chronic atrophic gastritis (CAG) based on serological markers and gastroscopic features.**

Methods: **A single-center cross-sectional study design was employed, consecutively enrolling 90 patients with pathologically confirmed CAG. Given the limited sample size (n = 90), the generalizability of the findings may be constrained. Clinical data, serological markers (pepsinogen I [PG I], pepsinogen II [PG II], pepsinogen I/II ratio [PGR], gastrin-17 [G-17]), and standardized gastroscopic features were collected. By using histopathologically confirmed high-grade dysplasia as the gold standard, patients were categorized into high-risk (n = 28) and low-risk (n = 62) groups. The Boruta algorithm was used to screen for important predictive variables, and three prediction models were constructed using multivariate logistic regression. Model performance was compared using area under the receiver operating characteristic curve (AUC), Akaike information criterion (AIC), and Bayesian information criterion (BIC).**

Results: **Univariate analysis revealed that patients in the high-risk group were older, had a significantly higher proportion with a family history of gastric cancer, and exhibited significantly lower PG I levels and PGR values. Additionally, they demonstrated significantly higher prevalence rates of open-type atrophy, moderate/severe intestinal metaplasia (IM), and suspicious lesions (all $p < 0.05$). The final model established by multivariate analysis (Model 3) integrated three core predictors: open-type atrophy (odds ratio [OR] = 9.51, 95% confidence interval [CI]: 2.77 - 32.67), moderate/severe IM (OR = 4.42, 95% CI: 1.30 - 15.02), and PG I (OR = 0.96, 95% CI: 0.93 - 0.98). This model demonstrated optimal predictive performance with an AUC of 0.895 (95% CI: 0.829 - 0.961) and the lowest AIC (76.2) and BIC (86.2) values.**

Conclusions: **This study successfully developed a simplified predictive model incorporating gastroscopic features and serum PG I levels, which effectively identifies high-grade dysplasia in patients with CAG. This model provides a potential practical tool for clinical risk stratification and precision management.**

(Clin. Lab. 2026;72:xx-xx. DOI: 10.7754/Clin.Lab.2025.251135)

Correspondence:
Li Zhang
Department of Traditional Chinese Medicine
Dingxi People's Hospital
No. 22, Anding Road
Anding District
Dingxi City
Gansu Province, 743000
China
Email: zhanglli66@outlook.com

Manuscript accepted December 19, 2025

KEYWORDS

chronic atrophic gastritis, high-grade intraepithelial neoplasia, risk prediction model, pepsinogen I, gastroscopic features

INTRODUCTION

Gastric cancer is a common malignant tumor worldwide and one of the leading causes of cancer-related mortality [1]. Its development is a multi-step progressive pro-

cess, with chronic atrophic gastritis (CAG) recognized as the most significant precancerous lesion for intestinal-type gastric cancer [2]. According to the classic "Correa sequence" model, the gastric mucosa progresses from a normal state through chronic inflammation, atrophy, intestinal metaplasia (IM), and dysplasia, ultimately potentially advancing to invasive carcinoma [3]. Within this process, high-grade dysplasia is regarded as the stage closest to malignancy [4], representing a critical window for clinical intervention and cancer prevention.

Accurate identification of patients with CAG at high risk for carcinogenic transformation is crucial for achieving early warning and precise management. Currently, gastroscopy and pathological biopsy serve as the "gold standard" for diagnosis and staging, with core assessment criteria including the extent and severity of gastric mucosal atrophy, the degree of intraepithelial neoplasia, and the presence of carcinoma in situ [5,6]. However, as an invasive procedure, gastroscopy is impractical for frequent, large-scale screening of all patients. In recent years, serological markers such as pepsinogen I (PG I), the pepsinogen I/II ratio (PGR), and gastrin-17 (G-17) have demonstrated significant potential as non-invasive biomarkers for assessing gastric mucosal status, particularly atrophy. Numerous studies have confirmed the significant value of these markers in detecting and evaluating gastric mucosal lesions. For instance, decreased levels of PG I and PGR are closely associated with gastric atrophy, and these markers can effectively predict the occurrence of gastric atrophy [7]. Furthermore, G-17 level is also considered an important indicator for evaluating gastric mucosal lesions, particularly demonstrating potential application value in the early screening of gastric cancer [8].

Although gastroscopic features and serological markers provide crucial information from morphological and functional perspectives respectively, their respective values in risk prediction differ. A key unmet need is how to effectively integrate information from these two sources to construct a simple, clinically applicable risk prediction tool that can more accurately identify CAG patients who have already developed or are at high risk of developing high-grade lesions. Such models could help optimize endoscopic surveillance strategies, enabling closer follow-up for high-risk patients while avoiding overtreatment of low-risk individuals.

Therefore, this study aimed to develop a simplified predictive model for assessing the risk of carcinogenic transformation (as indicated by high-grade intraepithelial neoplasia) in patients with CAG. This model was based on serological markers and standardized gastroscopic features, utilizing rigorous variable selection and model-building methodologies to provide practical guidance for clinical decision-making.

MATERIALS AND METHODS

Study population

This study employed a single-center, cross-sectional observational design. Between January 2023 and June 2025, eligible patients underwent simultaneous serological testing, gastroscopy, and pathological evaluation at a single time point. The study population comprised consecutively enrolled patients from the gastroenterology outpatient clinic or inpatient ward during this period. These patients underwent gastroscopy due to suspected CAG or routine gastric screening and strictly met the following inclusion criteria: 18 - 75 years; confirmed diagnosis of CAG via endoscopy combined with pathological biopsy (regardless of lesion location or extent); voluntary participation in the study with signed written informed consent. Exclusion criteria were as follows: a history of gastric surgery (e.g., subtotal gastrectomy), proton pump inhibitors, H2 receptor antagonists, or antibiotics within 2 weeks prior to examination (as these medications may significantly interfere with serological indicators); pathologically or radiologically confirmed gastric cancer; severe heart failure, liver failure, kidney failure, or hematological disorders; and female patients who were pregnant or lactating.

This study protocol was approved by the Ethics Review Committee of Dingxi People's Hospital. Based on the abovementioned criteria, a total of 90 patients were ultimately enrolled as subjects for this study.

Data collection

This study collected relevant data on the day patients that underwent gastroscopy through the hospital's electronic medical record system, laboratory information system, and structured data collection form system.

Clinical data: Demographic information was recorded first, including patient age (years) and gender; second, lifestyle and medical history data were collected, where smoking history was defined as continuous or cumulative smoking duration ≥ 6 months, and drinking history was defined as average daily alcohol intake > 30 g sustained for ≥ 1 year. Patient's primary symptoms at visit (e.g., presence of upper abdominal pain, postprandial bloating, early satiety), common comorbid chronic conditions (e.g., hypertension, type 2 diabetes), and long-term non-exclusion medications (e.g., nonsteroidal anti-inflammatory drugs, aspirin) were recorded. The current *Helicobacter pylori* infection was determined based on the results of the rapid urease test conducted during this endoscopy or the histopathological examination findings and was uniformly recorded as "positive" or "negative." Additionally, attention was paid to a family history of gastric cancer among the patient's first-degree relatives to comprehensively obtain the baseline clinical data required for the study.

Serological marker testing

Serum samples were collected from patients early on the morning of their gastroscopy. The procedure was as

follows: 5 mL of fasting blood were drawn from the patient's elbow vein. After standing at room temperature for 30 minutes, the sample was centrifuged at 3,000 rpm for 10 minutes. The supernatant serum was separated and stored at -80°C for subsequent testing. Serological indicator testing employed commercial ELISA or chemiluminescence assay kits, strictly adhering to kit instructions for detection of PG I, PG II, and G-17. All testing procedures were performed by a single laboratory physician unaware of patient group assignments to prevent detection bias.

Gastroscopic feature collection

Gastroscopy examinations were performed using high-definition electronic gastroscopy systems from Olympus or Pentax. Procedures and evaluations were conducted by two senior endoscopists with over five years of experience. To ensure consistency in assessment standards, all physicians received standardized training on gastroscopic feature evaluation criteria prior to the study. A pre-test was conducted: two physicians independently assessed a set of gastroscopy images (e.g., 20 cases) not included in the formal study. Inter-rater reliability (Kappa value) was calculated, with a Kappa value > 0.6 required before commencing the formal study. The examination procedure followed standardized protocols: All patients first underwent routine white-light gastroscopy. Afterward, electronic chromoendoscopy was conducted on suspicious lesions found during endoscopy and at standard sites according to the Sydney System protocol. Gastroscopic findings were documented in three aspects: First, the extent and classification of atrophy, recorded as “closed type” or “open type” based on the Kimura-Takemoto classification; second, the degree of IM, recorded as “none/mild” or “moderate/severe” based on the extent of white villous or patchy lesions observed during endoscopy; and third, macroscopically suspicious lesions, noting the presence of distinct nodules, erosions, ulcers, or polyps, along with the size and location of the lesions.

Outcome definition

The outcome was the presence of high-grade intraepithelial neoplasia, with pathological diagnosis serving as the gold standard. Biopsy procedures strictly adhered to the Sydney System protocol, with specific sampling sites including the gastric antrum (one tissue sample each from the greater curvature and lesser curvature), gastric angle (one tissue sample), and gastric body (one tissue sample each from the greater curvature and lesser curvature). Additionally, any macroscopically suspicious lesions identified during endoscopy underwent supplementary biopsy sampling. All pathological specimens obtained were independently reviewed and diagnosed by two senior pathologists unaware of the patients' clinical information or endoscopy findings. Diagnoses strictly adhered to the relevant classification criteria of the World Health Organization, categorizing patients into two groups: the low-risk group (CAG with or without low-grade intraepithelial neoplasia) and the high-risk group (high-grade intraepithelial neoplasia). Discrepancies were resolved by a third senior pathologist.

Statistical analysis

Statistical analysis was performed using SPSS 26.0 software and R software (version 4.2.1). All hypothesis tests were two-tailed, with $p < 0.05$ set as the criterion for statistically significant differences. In statistical analysis, continuous variables meeting normal distribution were expressed as mean ± standard deviation (SD) and analyzed using independent samples *t*-tests. Non-normally distributed variables were expressed as median (interquartile range, IQR) and analyzed using Mann-Whitney U tests. Categorical data were expressed as number (percentage, %) and compared between groups using the chi-squared test or Fisher's exact test.

Variance inflation factor (VIF) was employed to assess multicollinearity, with VIF < 10 serving as the criterion for no significant collinearity. Subsequently, the Boruta algorithm was employed to select the top 6 most important independent variables. A multivariate binary logistic regression model was constructed, incorporating quantitative parameters as continuous values and categorical variables as dummy variables. Three models were developed: one containing only gastroscopic features, one containing only demographic and serological indicators, and one integrating key indicators. The Wald test was used to calculate the odds ratio (OR) and its 95% confidence interval (CI). Model performance was evaluated based on goodness of fit and differentiation power: goodness of fit was compared using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), with lower values indicating superior models [9]; discriminatory power was assessed by the area under the receiver operating characteristic curve (AUC) and its 95% CI.

RESULTS

Comparison of baseline characteristics between high-risk and low-risk groups

This study included 90 patients with CAG, comprising 28 cases (31.1%) in the high-risk group (with high-grade intraepithelial neoplasia) and 62 cases (68.9%) in the low-risk group. As shown in Table 1, patients in the high-risk group were significantly older than those in the low-risk group (63 ± 7 years vs. 57 ± 8 years, $p < 0.001$) and had a higher proportion with a family history of gastric cancer (28.6% vs. 9.7%, $p = 0.030$). The prevalence of active *Helicobacter pylori* infection showed an increasing trend in the high-risk group (57.1% vs. 37.1%, $p = 0.076$). No significant differences were observed between the two groups in terms of gender, lifestyle, comorbidities, medication history, or clinical symptoms (all $p > 0.05$).

Table 1. Comparison of baseline characteristics between high-risk and low-risk groups.

Characteristic	Groups		p-value
	High-risk	Low-risk	
	n = 28	n = 62	
Demographics	63 ± 7	57 ± 8	< 0.001
Age, years			
Gender (male)	17 (60.7%)	30 (48.4%)	0.278
History and Lifestyle, n (%)			
Smoking	8 (28.6%)	16 (25.8%)	0.784
Drinking	7 (25.0%)	14 (22.6%)	0.802
Current H. pylori	16 (57.1%)	23 (37.1%)	0.076
Family history	8 (28.6%)	6 (9.7%)	0.03
Comorbidities and Medication, n (%)			
Hypertension	11 (39.3%)	19 (30.6%)	0.421
Type 2 diabetes	6 (21.4%)	10 (16.1%)	0.561
NSAIDs use	2 (7.1%)	9 (14.5%)	0.492
Aspirin use	5 (17.9%)	10 (16.1%)	1
Presenting symptoms			
Epigastric pain	16 (57.1%)	41 (66.1%)	0.413
Postprandial fullness	15 (53.6%)	26 (41.9%)	0.305
Early satiety	12 (42.9%)	19 (30.6%)	0.259

Normally distributed continuous data are presented as mean ± SD, compared by Student's *t*-test; categorical data are presented as n (%), compared by chi-squared test or Fisher's exact test.

Table 2. Comparison of laboratory indicators and gastroscopic features.

Characteristic	Groups		p-value
	High-risk	Low-risk	
	n = 28	n = 62	
Serological markers, n (%)			
Pg I, ng/mL	53 (41, 67)	78 (59, 106)	< 0.001
Pg II, ng/mL	17 (14, 24)	16 (12, 22)	0.398
PGR	3.5 (2.7, 5.0)	5.2 (3.3, 9.4)	0.005
G-17, pmol/L	21 (16, 28)	18 (13, 26)	0.091
Endoscopic features, n (%)			
Open atrophy	20 (71.4%)	14 (22.6%)	< 0.001
Moderate/Severe IM	21 (75.0%)	23 (37.1%)	< 0.001
Suspicious lesions	11 (39.3%)	9 (14.5%)	0.009

Non-normally distributed continuous data are presented as median [IQR], compared by Mann-Whitney U test; categorical data are presented as n (%), compared by chi-squared test or Fisher's exact test. GC: gastric cancer, PgI: pepsinogen I, PgII: pepsinogen II, PGR: pepsinogen I/II ratio, G-17: gastrin-17, moderate/severe IM: moderate/severe intestinal metaplasia.

Table 3. Construction of a risk model for the carcinogenic transformation of chronic atrophic gastritis.

	Model 1			Model 2			Model 3		
Characteristic	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**	**OR**	**95% CI**	**p-value**
Open atrophy	**7.63**	**2.54, 22.94**	**< 0.001**				**9.51**	**2.77, 32.62**	**< 0.001**
Moderate/Severe IM	**4.17**	**1.37, 12.73**	**0.012**				**4.42**	**1.30, 15.04**	**0.017**
Suspicious lesions	**3.47**	**0.95, 12.71**	**0.06**						
Age, years				**1.13**	**1.04, 1.22**	**0.003**			
PGR				**0.84**	**0.65, 1.09**	**0.19**			
PgI, ng/mL				**0.98**	**0.95, 1.01**	**0.192**	**0.96**	**0.93, 0.98**	**0.001**

CI: confidence interval, OR: odds ratio.
Model 1: Null df = 89; Log-likelihood = -40.2; AIC = 88.4; BIC = 98.4; deviance = 80.4.
Model 2: Null df = 89; Log-likelihood = -40.8; AIC = 89.6; BIC = 99.6; deviance = 81.6.
Model 3: Null df = 89; Log-likelihood = -34.1; AIC = 76.2; BIC = 86.2; deviance = 68.2.

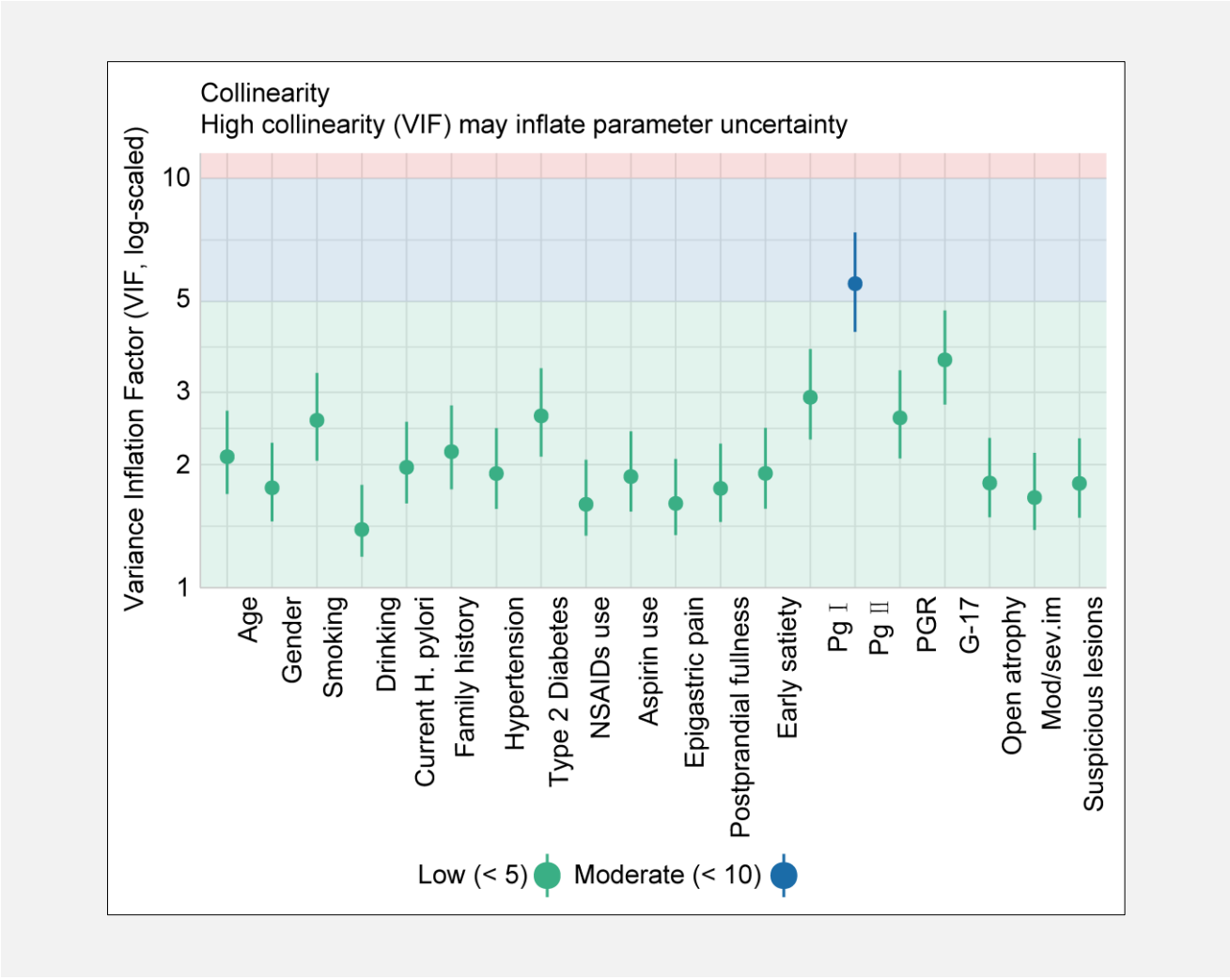


Figure 1. VIF of predictor variables in the prediction model.

This figure displays the VIF values and their ranges for all predictor variables. All VIF values are well below the commonly accepted threshold of 10, indicating no severe multicollinearity among the variables and satisfying the prerequisites for multivariate logistic regression analysis.

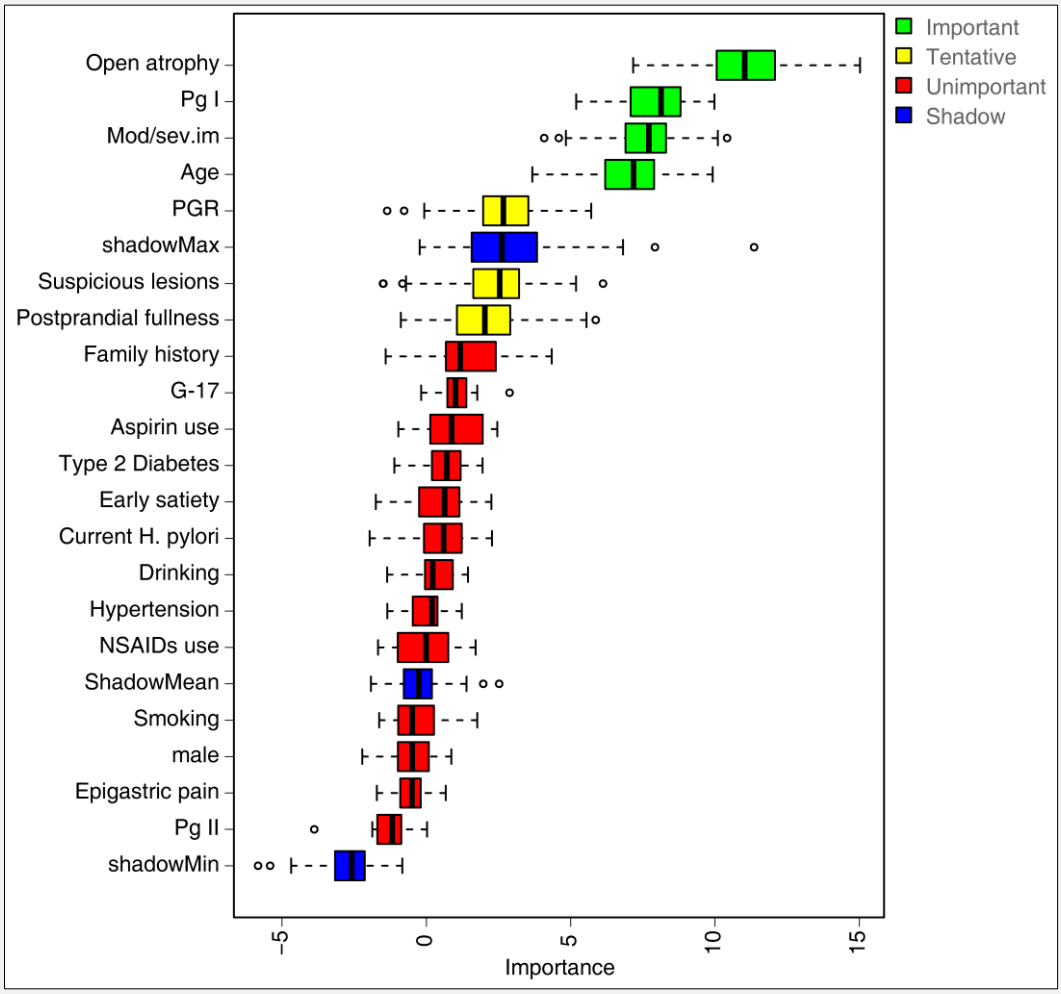


Figure 2. Prediction variable importance ranking based on the Boruta algorithm.

Variable importance screening was performed using the Boruta algorithm based on random forests. Ultimately, the top six variables by importance (including "open-type atrophy" and "PG I") were selected for subsequent model construction.

Comparison of laboratory markers and gastroscopic features

As shown in Table 2, regarding serological markers, patients in the high-risk group exhibited significantly lower PG I levels ($p < 0.001$) and markedly reduced PGR ($p = 0.005$) compared to the low-risk group. No significant differences were observed between the two groups in PG II and G-17 levels ($p > 0.05$).

Regarding gastroscopic features, the high-risk group exhibited significantly higher proportions of open-type atrophy, moderate/severe IM, and macroscopically suspicious lesions compared to the low-risk group, with highly statistically significant differences (71.4% vs. 22.6%, $p < 0.001$; 75.0% vs. 37.1%, $p < 0.001$; 39.3% vs. 14.5%, $p = 0.009$).

Construction and performance evaluation of a risk model for carcinogenic transformation in CAG

Under rigorous statistical assumptions, multicollinearity diagnostics were performed for all independent variables. Results showed that the VIFs for all variables were below 10 (Figure 1), indicating that multicollinearity did not affect subsequent model construction. Subsequently, the Boruta algorithm (an improved variable selection method based on Random Forests) was employed for feature importance assessment. The results are shown in Figure 2, based on which the top 6 important predictor variables were selected for modeling. Model 1 included only gastroscopic features. Results showed that open-type atrophy (OR = 7.63, 95% CI: 2.54 - 22.93, $p < 0.001$) and moderate/severe IM (OR = 4.17,

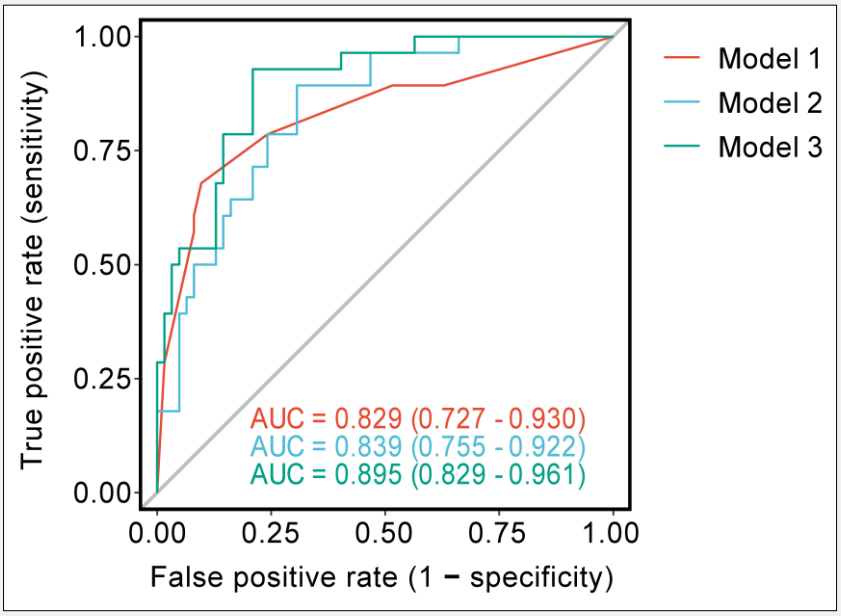


Figure 3. ROC curves for three predictive models.

The discriminatory capabilities of three predictive models for CAG cancer risk were compred. AUC and its 95% CI for Model 1 (gastroscopic features only), Model 2 (age and serological indicators), and Model 3 (integrated model) were analyzed.

95% CI: 1.37 - 12.66, p = 0.012) were independent significant predictors of carcinogenic risk, while the presence of suspicious lesions was not a statistically significant predictor (p = 0.060). Model 2 incorporated demographic and serological indicators, with only age (OR = 1.13, 95% CI: 1.04 - 1.22, p = 0.003) emerging as a significant predictor, while PGR and PG I were non-significant (p > 0.05). Model 3 demonstrated the best overall performance, integrating the top three most significant variables (open-type atrophy, PG I, and moderate/severe IM). Open-type atrophy (OR = 9.51, 95% CI: 2.77 - 32.67, p < 0.001) and moderate/severe IM (OR = 4.42, 95% CI: 1.30 - 15.02, p = 0.017) remained highly significant, while serum PG I level emerged as a significant negative predictor (OR = 0.96, 95% CI: 0.93 - 0.98, p < 0.001). Model 3 possessed the lowest AIC value (76.2) and BIC value (86.2), significantly outperforming Model 1 (AIC = 88.4, BIC = 98.4) and Model 2 (AIC = 89.6, BIC = 99.6), suggesting optimal model fit and interpretability. Therefore, Model 3, which integrates gastroscopic features with the key serological indicator (PG I), represents the optimal simplified model for predicting the risk of carcinogenic transformation in CAG. Additionally, the differentiation capabilities of the three models were evaluated using ROC curves (Figure 3). The AUC for Models 1, 2, and 3 were 0.829 (95% CI: 0.727 - 0.930), 0.839 (95% CI: 0.755 - 0.922), and 0.895 (95% CI: 0.829 - 0.961), respectively. Model 3 exhibited a significantly higher AUC value than the other two models, indicating its optimal predictive discriminatory capability for the risk of carcinogenic transformation in CAG.

DISCUSSION

This study successfully constructed and validated a simplified predictive model for identifying CAG patients with high-grade intraepithelial neoplasia through a standardized variable screening and modeling process. The core finding is that Model 3 - which integrates key gastroscopic features with serological markers (including open-type atrophy, moderate/severe IM, and PG I) - demonstrated optimal predictive performance (AUC = 0.895), outperforming models relying solely on gastroscopic features or serological markers. This result confirms that combining morphological and functional information enables more effective stratification of cancer risk in CAG patients.

This study identified open-type atrophy, moderate/severe IM, and serum PG I as core predictive variables through the Boruta algorithm. Model 3 (open-type atrophy: OR = 9.51, 95% CI: 2.77 - 32.67, p < 0.001; moderate/severe IM: OR = 4.42, 95% CI: 1.30 - 15.02, p =

0.017; PG I: OR = 0.96, 95% CI: 0.93 - 0.98, p < 0.001) demonstrated optimal predictive performance (AIC = 76.2, BIC = 86.2, AUC = 0.895). This finding strongly aligns with the "Correa sequence" theory [10], which posits that the progression from atrophy and IM to dysplasia represents the core pathway of carcinogenesis. Open-type atrophy (Kimura-Takemoto classification), characterized by atrophy extending beyond the gastric angle into the body, indicates more extensive and irreversible mucosal damage [11]. Its status as a strong predictor (highest OR value) aligns with prior research findings-patients with open-type atrophy face a heightened risk of progression to high-grade neoplasia [12]. Furthermore, severe IM in this study increased the risk by 4.4-fold. In literature supporting Correa's sequence theory, metaplasia represents a process of cellular replacement triggered by environmental stimuli, serving as a precursor to low-grade dysplasia and potentially progressing to high-grade dysplasia and cancer [11]. Moreover, the occurrence of metaplasia may be associated with transcription factors and epigenetic reprogramming [13], offering new insights into the molecular mechanisms underlying gastric mucosal lesions.

PG I, as a functional marker of gastric body atrophy [14], demonstrated independent predictive value in Model 3. Even after adjusting for gastroscopic atrophy classification, PG I levels remained significantly associated with high-grade lesion risk (OR = 0.96). This value indicates that a 1 ng/mL decrease in PG I correlates with a 4% increase in high-grade neoplastic risk. Its negative predictive effect further validates its intrinsic function: PG I is primarily secreted by chief cells of the fundic glands, and reduced levels directly reflect the degree of gastric body gland atrophy. Such atrophy is accompanied by decreased gastric acid secretion, leading to gut microbiota imbalance (e.g., increased nitrite-reducing bacteria), accelerated carcinogen production, and ultimately heightened cancer risk [15,16].

This mechanism also explains the unique value of PG I - the information it provides on "functional atrophy" compensates for the limitations of "morphological atrophy" assessment under gastroscopy, with the two being complementary. Furthermore, the findings regarding PG I in this study resonate with those from multiple other investigations. Research indicates that individuals with abnormal serum PG levels exhibit a significantly higher cumulative incidence of gastric cancer compared to normal controls, particularly among those concurrently presenting abnormal PG levels and *H. pylori* positivity, who demonstrate the highest gastric cancer incidence rate (56.0) [17]. Another study highlights that low PG I levels correlate not only with gastric cancer development but also with the occurrence of multiple gastric epithelial neoplasms (GENs). Research indicates that among patients undergoing endoscopic resection, PG I levels below 30 ng/mL are significantly associated with the occurrence of multiple GENs [18]. Collectively, these studies provide crucial theoretical support for "combining endoscopy with serological markers for risk stratification of gastric precancerous lesions" and validate the necessity of serological indicators in risk assessment.

The predictive performance of the standalone gastroscopy feature model (Model 1, AUC = 0.829) or the demographic-serological model (Model 2, AUC = 0.839) both demonstrated lower predictive efficacy than the integrated model (Model 3, AUC = 0.895). Model 3 exhibited the lowest AIC and BIC values, indicating that integrating "morphological (gastroscopic features) + functional (serum PG I)" information significantly enhances predictive accuracy. This advantage stems from the following: gastroscopic features (e.g., open-type atrophy, moderate/severe IM) visually reflect the "spatial extent and pathological stage" of mucosal lesions but cannot quantify the degree of mucosal functional impairment [19]. In contrast, serum PG I enables non-invasive assessment of gastric gland function, compensating for endoscopy's limitations in evaluating "functional aspects" [20]. For example, some patients exhibit only mild atrophy on gastroscopy but show significantly reduced PG I levels, indicating impaired gastric body gland function and higher potential carcinogenic risk. Relying solely on gastroscopy for assessment may misclassify such patients as low-risk, whereas the integrated model effectively avoids this issue. Previous studies have only demonstrated the role of these serum markers in predicting early gastric cancer but have not achieved integration [21]. Our findings suggest that gastroscopy and serological testing exhibit synergistic effects in CAG risk management, potentially offering novel insights for developing more efficient risk stratification strategies.

Analysis of several traditional risk factors in this study also provided valuable insights. In a previous multicenter cohort study, researchers found that *H. pylori* infection is associated with a significantly increased risk of gastric cancer, particularly non-cardia gastric cancer and cardia gastric cancer [22]. In this study, although the prevalence of current *H. pylori* infection showed an upward trend in the high-risk group (57.1% vs. 37.1%), it did not reach statistical significance. This phenomenon may reflect the "dual role" of *H. pylori* infection in gastric cancer development: while *H. pylori* serve as initial driver of CAG, their bacterial load may decrease as gastric microenvironment changes during disease progression to advanced stages. Consequently, the prevalence of active infection carries reduced weight when assessing the immediate carcinogenic risk in patients with established CAG.

Additionally, it is noteworthy that no significant differences were observed between the two groups in terms of dyspepsia symptoms and common comorbidities. Symptoms such as epigastric pain and postprandial bloating showed no correlation with the severity of gastric mucosal lesions, potentially reflecting their low specificity and high overlap between patients with atrophic gastritis and functional dyspepsia. Similarly, systemic metabolic factors like hypertension and diabetes

did not demonstrate an association with carcinogenic risk, suggesting these factors may exert weaker influence on local gastric mucosal carcinogenesis than the progression of mucosal lesions themselves.

To optimize stratified management strategies for CAG patients, addressing the current clinical practice of "uniform frequency" follow-ups (e.g., annual gastroscopy) often leading to under-monitoring of high-risk patients and overtreatment of low-risk patients, Model 3 developed in this study incorporates only three readily accessible indicators: endoscopic atrophy classification, degree of IM, and serum PG I. Clinicians can rapidly calculate risk scores and implement tiered patient management. High-risk patients (e.g., those with open-type atrophy + moderate/severe IM + PG I < 70 ng/mL) should undergo endoscopic follow-up every 3 - 6 months, with focused monitoring of suspicious lesions to enable timely intervention for high-grade neoplasia. Low-risk patients (closed-type atrophy + mild IM + normal PG I) may extend follow-up intervals to 1 - 2 years, reducing endoscopic burden. This strategy aligns closely with recommendations for "risk-adjusted follow-up frequency" [23] and significantly enhances clinical management efficiency. Meanwhile, considering that the invasiveness and cost of gastroscopy limit its application in primary care hospitals and large-scale screening programs, while serum PG I testing is convenient and relatively inexpensive [24], this model can first use serum PG I to identify high-risk individuals for "PG I reduction". High-risk individuals with reduced PG I should receive endoscopy to substantially reduce unnecessary procedures. For instance, in community screening, serum PG I could be tested in individuals aged 50 and above, with those showing PG I < 70 ng/mL then undergoing endoscopy. This approach enhances screening efficiency while lowering healthcare costs, aligning with the policy direction of "tiered diagnosis and treatment" [25].

This study has several limitations that warrant clarification. First, as a single-center cross-sectional study with a sample size of only 90 cases (28 in the high-risk group), all patients were enrolled from the same hospital. Their characteristics, such as *H. pylori* infection rates and dietary habits, may differ from those of populations in other regions. This design may introduce selection bias and limits model generalizability. Additionally, the cross-sectional approach precludes observation of temporal relationships between predictor variables and outcomes (e.g., whether PG I reduction precedes high-grade neoplasia). Future studies should employ a multicenter prospective cohort with ≥ 300 participants to validate the model's predictive efficacy across diverse populations and assess its long-term predictive value at 1 and 3 years. Second, the relatively limited sample size may affect model stability despite the use of Bootstrap internal validation. Third, the single-center design may introduce selection bias, limiting the generalizability of results. Furthermore, despite collecting extensive clinical variables, unmeasured confounding factors may still exist.

CONCLUSION

In summary, this study developed a simplified predictive model incorporating open-type atrophy, moderate/severe IM, and PG I, demonstrating good performance in distinguishing CAG patients with high-grade intraepithelial neoplasia. Future research should focus on external validation in prospective, multicenter cohorts and explore the feasibility of translating this model into a clinical scoring system, ultimately providing a practical tool for achieving precise risk management in CAG patients.

Declaration of Generative AI in Scientific Writing:
During the preparation of this work, the authors did not use generative artificial intelligence or AI-assisted technologies.

Availability of Data and Materials:
The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Ethical Approval and Consent to Participate:
The present study was approved by the Ethics Committee of Dingxi People's Hospital, and written informed consent was provided by all patients prior to the study start. All procedures were performed in accordance with the ethical standards of the Institutional Review Board and the Declaration of Helsinki and its later amendments or comparable ethical standards.

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

References:

1. Vannella L, Sbrozzi-Vanni A, Lahner E, et al. Development of type I gastric carcinoid in patients with chronic atrophic gastritis. Aliment Pharmacol Ther 2011;33(12):1361-9. (PMID: 21492197)

2. Watari J, Chen N, Amenta PS, et al. Helicobacter pylori associated chronic gastritis, clinical syndromes, precancerous lesions, and pathogenesis of gastric cancer development. World J Gastroenterol 2014;20(18):5461-73. (PMID: 24833876)

3. Li F, Wang Y, Ping X, et al. Molecular evolution of intestinal-type early gastric cancer according to Correa cascade. J Biomed Res 2024;39(3):270-85. (PMID: 39314047)

4. Fassan M, Simbolo M, Bria E, et al. High-throughput mutation profiling identifies novel molecular dysregulation in high-grade intraepithelial neoplasia and early gastric cancers. Gastric Cancer 2014;17(3):442-9. (PMID: 24272205)

5. Minoda Y, Chinen T, Osoegawa T, et al. Superiority of mucosal incision-assisted biopsy over ultrasound-guided fine needle aspiration biopsy in diagnosing small gastric subepithelial lesions: a propensity score matching analysis. BMC Gastroenterol 2020; 20(1):19. (PMID: 31964357)

6. Wang Y, Zhou J, Meng N, et al. The occurrence, progression and development of four types of gastric mucosal atrophic lesions and their histopathological characteristics. Gastric Cancer 2023;26(5): 721-33. (PMID: 37328675)

7. Song HJ, Jang SJ, Yun SC, et al. Low Levels of Pepsinogen I and Pepsinogen I/II Ratio are Valuable Serologic Markers for Predicting Extensive Gastric Corpus Atrophy in Patients Undergoing Endoscopic Mucosectomy. Gut Liver 2010;4(4):475-80. (PMID: 21253295)

8. Yu G, Wang GX, Wang HG, Mo FF, Tang BB. The value of detecting pepsinogen and gastrin-17 levels in serum for pre-cancerous lesion screening in gastric cancer. Neoplasma 2019;66(4): 637-40. (PMID: 31058531)

9. Glatting G, Kletting P, Reske SN, Hohl K, Ring C. Choosing the optimal fit function: comparison of the Akaike information criterion and the F-test. Med Phys 2007;34(11):4285-92. (PMID: 18072493)

10. Al-Awadhi H, John R, Al-Marzooqi F, Vincze A, Branicki F, Karam SM. Sequential alterations in gastric biopsies and tumor tissues support the multistep process of carcinogenesis. Histol Histopathol 2011;26(9):1153-64. (PMID: 21751147)

11. Giroux V, Rustgi AK. Metaplasia: tissue injury adaptation and a precursor to the dysplasia-cancer sequence. Nat Rev Cancer 2017;17(10):594-604. (PMID: 28860646)

12. Shichijo S, Hirata Y, Niikura R, et al. Histologic intestinal metaplasia and endoscopic atrophy are predictors of gastric cancer development after Helicobacter pylori eradication. Gastrointest Endosc 2016;84(4):618-24. (PMID: 26995689)

13. Jiang H, Li G. Transcription factors direct epigenetic reprogramming at specific loci in human cancers. Front Genet 2023;14: 1234515. (PMID: 37876590)

14. Mansour-Ghanaei F, Joukar F, Baghaee M, Sepehrimanesh M, Hojati A. Only serum pepsinogen I and pepsinogen I/II ratio are specific and sensitive biomarkers for screening of gastric cancer. Biomol Concepts 2019;10(1):82-90. (PMID: 31188744)

15. Sui Z, Chen J, Li P, et al. Risk for gastric cancer in patients with gastric atrophy: a systematic review and meta-analysis. Transl Cancer Res 2020;9(3):1618-24. (PMID: 35117509)

16. Wang Z, Gao X, Zeng R, et al. Changes of the Gastric Mucosal Microbiome Associated With Histological Stages of Gastric Carcinogenesis. Front Microbiol 2020;11:997. (PMID: 32547510)

17. Zhang X, Xue L, Xing L, et al. Low serum pepsinogen I and pepsinogen I/II ratio and Helicobacter pylori infection are associated with increased risk of gastric cancer: 14-year follow up result in a rural Chinese community. Int J Cancer 2012;130(7):1614-9. (PMID: 21547904)

18. Park SY, Lim SO, Ki HS, et al. Low pepsinogen I level predicts multiple gastric epithelial neoplasias for endoscopic resection. Gut Liver 2014;8(3):277-81. (PMID: 24827624)

19. Tziatzios G, Ziogas D, Gkolfakis P, et al. Endoscopic Grading and Sampling of Gastric Precancerous Lesions: A Comprehensive Literature Review. Curr Oncol 2024;31(7):3923-38. (PMID: 39057162)

20. Wasielica-Berger J, Rogalski P, Pryczynicz A, Swidnicka-Siergiejko A, Dabrowski A. Methylene blue chromoendoscopy is more useful in detection of intestinal metaplasia in the stomach than mucosal pit pattern or vessel evaluation and predicts advanced Operative Link on Gastric Intestinal Metaplasia stages. Clin Endosc 2023;56(2):203-13. (PMID: 36890637)

21. Wang Y, Zhu Z, Liu Z, et al. Diagnostic value of serum pepsinogen I, pepsinogen II, and gastrin-17 levels for population-based screening for early-stage gastric cancer. J Int Med Res 2020; 48(3):300060520914826. (PMID: 32228342)

22. Yan X, Zeng H, Li H, et al. The current infection with Helicobacter pylori and association with upper gastrointestinal lesions and risk of upper gastrointestinal cancer: Insights from multicenter population-based cohort study. Int J Cancer 2024;155(7):1203-11. (PMID: 38712628)

23. Ciechanski K, Ciechanski E, Klosowska-Kapica K, Skrzydlo-Radomanska B. Single-Time Gastroscopy in High-Risk Patients: Screening Effectiveness for Gastric Precancerous Conditions in a Low-To Moderate-Incidence Population. J Clin Med 2025; 14(19):6910. (PMID: 41095991)

24. Enriquez-Sanchez LB, Gallegos-Portillo LG, Camarillo-Cisneros J, et al. Cost-benefit of serum pepsinogen screening for gastric adenocarcinoma in the Mexican population. Rev Gastroenterol Mex (Engl Ed) 2022;87(3):285-91. (PMID: 34794924)

25. Zhao Z, Xie X, Wu Q. Medical staff evaluation on 'the effect of medical alliance policy on hierarchical diagnosis and treatment'. Front Public Health 2024;12:1366100. (PMID: 38813406)