Lan *et al. BMC Cancer* (2025) 25:518
https://doi.org/10.1186/s12885-025-13860-3

BMC Cancer

RESEARCH

Open Access

# A risk prediction model for gastric cancer based on endoscopic atrophy classification

Yadi Lan[1], Weijia Sun[1], Shen Zhong[1], Qianqian Xu[1], Yining Xue[2], Zhaoyu Liu[1], Lei Shi[2], Bing Han[1], Tianyu Zhai[1], Mingyue Liu[1], Yujing Sun[1] and Hongwei Xu[1,2*]

## Abstract

**Backgrounds** Gastric cancer (GC) is a prevalent malignancy affecting the digestive system. We aimed to develop a risk prediction model based on endoscopic atrophy classification for GC.

**Methods** We retrospectively collected the data from January 2020 to October 2021 in our hospital and randomly divided the patients into training and validation sets in an 8:2 ratio. We used multiple machine learning algorithms such as logistic regression (LR), Decision tree, Support Vector Machine, Random forest, and so on to establish the models. We employed the Least absolute shrinkage and selection operator (LASSO) to screen variables for the LR model. However, we chose all the variables to construct the models for other machine learning algorithms. All models were evaluated using the receiver operating characteristic curve (ROC), predictive histograms, and decision curve analysis (DCA).

**Results** A total of 1156 patients were selected for the analysis. Five variables, including age, sex, family history of GC, HP infection status, and Kimura-Takemoto Classification (KTC), were screened using LASSO analysis. The area under the curve (AUC) of all the machine learning models ranged from 0.762 to 0.974 in the training set and from 0.608 to 0.812 in the validation set. Among them, the LR model exhibited the highest AUC value (0.812, 95%CI: 0.737–0.887) in the validation set with good calibration and clinical applicability. Finally, we constructed a nomogram to demonstrate the LR model.

**Conclusions** We established a nomogram based on endoscopic atrophy classification for GC, which might be valuable in predicting GC risk and assisting clinical decision-making.

**Keywords** Endoscopy, Gastric cancer, Machine learning, Model, Prediction

*Correspondence:
Hongwei Xu
xhwsdslyy@sina.com
[1] Department of Gastroenterology, Shandong Provincial Hospital, Shandong University, Jinan, Shandong 250021, China
[2] Department of Gastroenterology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong 250021, China

## Introduction

Gastric cancer (GC) is a prevalent malignancy affecting the digestive system. The Global Cancer Statistics report for 2020 revealed that more than one million people worldwide have been diagnosed with GC, which accounts for 5.6% of all cancers [1]. Unfortunately, the five-year survival rate for progressive GC remains poor, with a rate of less than 10% [2]. However, patients with early GC have a significantly better prognosis, with survival rates surpassing 90% after endoscopic treatment [3]. Consequently, the development of risk prediction models for GC is crucial. Accurate risk prediction enables clinicians to recommend appropriate screening strategies, surveillance programs, and personalized treatment plans.

Gastric mucosal atrophy, a precancerous condition of the stomach, is a crucial stage in the development of GC [4]. The most widely accepted pattern is Correa's cascade [5]. Many studies have demonstrated that the risk of gastric carcinogenesis was associated with the degree and extent of atrophy [6]. Currently, the updated Sydney System [7] and the operative link on gastritis assessment (OLGA) [8] are mostly used to assess the degree and extent of gastric atrophy. However, these two systems require biopsies from five sites under endoscopy, which increases patients' medical costs. In addition, it appears that a biopsy is not sufficient to evaluate the condition of the entire gastric mucosa. Besides histopathology, various serological indicators such as pepsinogen, Helicobacter pylori (H. pylori) antibodies, gastrin-17, and tumor markers could help assess gastric mucosa atrophy or predict the risk of GC [9]. However, these indicators are susceptible to many factors, such as the testing methods, the use of anti-acid medications, and the comorbidity of other gastric diseases.

The Kimura-Takemoto classification (KTC), proposed by Japanese scholars in 1969 [10], is primarily used to assess the extent of gastric mucosal atrophy during endoscopy. Unlike serological indicators, the KTC is not susceptible to other factors and does not undergo significant alteration in the short term. In addition, the international uniform standard of KTC helps increase the applicability and reliability of the model. Several studies [11, 12] have revealed high diagnostic concordance between the KTC and histopathology, making it a promising tool for identifying 'high-risk' endoscopic screening individuals.

Artificial intelligence is developing rapidly, and its integration with the field of medicine is increasing. Machine learning plays a crucial role in this integration [13]. Machine learning analyzes data from multiple dimensions and continuously learns from the data to improve algorithms, which makes it particularly useful for disposing of complex medical data and generating personalized risk assessments based on individual profiles [14]. Many researchers have already used machine learning techniques for early cancer prediction and have achieved some success.

Currently, there is no feasible and efficient endoscopic risk prediction model for GC in China. We aimed to incorporate various machine learning algorithms to establish a useful GC risk prediction model based on the endoscopic atrophy classification, which could guide clinical decision-making.

## Methods

### Study population

It is a retrospective cross-sectional study. Patients who underwent endoscopic examination for gastrointestinal symptoms at our Hospital from January 2020 to October 2021 were consecutively selected. The inclusion criteria were as follows: (1) aged between 30 and 90 years; (2) underwent endoscopic assessment of gastric atrophy; (3) had complete medical records. The exclusion criteria were: (1) esophageal cancer; (2) history of gastric surgery.

### Data collection

We collected the patients' information from the electronic medical records and endoscopic reports, including age, sex, family history of GC in first-degree relatives, smoking, alcohol, and Kimura Takemoto Classification (KTC). We classified the H.pylori infection status into two groups - uninfected and infected (whether current or post-eradication). The diagnostic criteria for Helicobacter pylori (HP) are as following [15]: (1) positive C13 breath test; (2) positive HP antibodies; (3) presence of HP in pathological biopsy samples.

Two endoscopists assessed the KTC for all patients according to the endoscopic images. Both of the two endoscopists have more than ten years of experience. The criteria for KTC [16] were as follows: (1) C1, the atrophy confined in the antrum; (2) C2, the atrophy exceeded the incisura angularis but confined in the lesser curvature; (3) C3, the atrophy was in the lesser curvature and did not exceed the cardia; (4) O1, the atrophy extended to the cardia and the atrophic border was between the lesser curvature and the anterior wall; (5) O2, the atrophic border was in the anterior wall; (6) O3, the atrophic border was between the anterior wall and the greater curvature; C0 meant those without atrophy. Endoscopic manifestations of atrophy mainly included the appearance of the capillary network, pallor of the gastric mucosa, and flattening or even the absence of the mucosal folds. The definition of GC included high-grade intraepithelial neoplasia and invasive carcinoma of the stomach, and the diagnostic criteria referred to the Vienna classification for gastrointestinal epithelial neoplasia [17, 18].

**Table 1** The baseline of the patients

| | Total | Non-GC | GC | *P* |
|---|---|---|---|---|
| Total | 1156 | 1006 | 150 | |
| **sex**, **n(%)** | | | | < 0.001 |
| female | 487(42.1) | 444(44.1) | 43(28.7) | |
| male | 669(57.9) | 562(55.9) | 107(71.3) | |
| **age** | | | | < 0.001 |
| mean (SD) | 58(12) | 57(12) | 65(8) | |
| **smoking**, **n(%)** | | | | 0.021 |
| NO | 817(70.7) | 723(71.9) | 94(62.7) | |
| YES | 339(29.3) | 283(28.1) | 56(37.3) | |
| **alcohol**, **n(%)** | | | | 0.089 |
| NO | 720(62.3) | 636(63.2) | 84(56.0) | |
| YES | 436(37.7) | 370(36.8) | 66(44.0) | |
| **FamilyHistory**, **n(%)** | | | | < 0.001 |
| NO | 1117(96.6) | 980(97.4) | 137(91.3) | |
| YES | 39(3.4) | 26(2.6) | 13(8.7) | |
| **KTC**, **n(%)** | | | | < 0.001 |
| C0-C1 | 839(72.6) | 765(76.0) | 74(49.3) | |
| C2-C3 | 252(21.8) | 197(19.6) | 55(36.7) | |
| O1-O3 | 65(5.6) | 44(4.4) | 21(14.0) | |
| **HP**, **n(%)** | | | | 0.369 |
| NO | 579(50.1) | 509(50.6) | 70(46.7) | |
| YES | 577(49.9) | 497(49.4) | 80(53.3) | |

### Statistical analysis

We employed IBM SPSS Statistics (Windows, version 26.0) and Python (version 3.0) software for the statistical analysis. We used mean and standard deviation (SD) to describe continuous variables and frequencies with percentages for categorical variables. There were no missing data in our study. We randomly divided the data into the training and validation set in an 8:2 ratio. The same patients were not both in the training and validation set. We employed multiple machine learning models such as logistic regression (LR), NaiveBayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Adaptive Boosting (AdaBoost), Random forest (RF), Decision Tree, ExtraTrees, XGBoost, LightGBM, and gradient boosting algorithms to construct the models. We used the Least Absolute Shrinkage and Selection Operator (LASSO) to select variables for the LR model. However, we chose all the variables without screening for the other models. We used the area under the receiver operating characteristic curve (AUC) to evaluate the models' discrimination, prediction histograms to assess the calibration, and decision curve analysis (DCA) to evaluate clinical utility. Finally, we constructed the nomogram to demonstrate the LR model.

## Results

### The baseline of the patients

In this study, we finally selected 1156 patients for data analysis, including 150 patients with GC and 1006 non-GC patients. The flowchart for patient selection is in Figure S1. The mean age of the patients was 58 ± 12 years, including 669 (57.87%) males, 339 (29.33%) smokers, 436 (47.72%) alcohol drinkers, and 39 (3.37%) with a family history of GC. The detailed information is in Table 1.

### Variable selection

We randomly divided the data into training and validation sets, with 924 patients in the training set and 232 patients in the validation set. Five variables were selected using LASSO analysis, including sex, age, family history of GC, HP infection status, and KTC (Fig. 1), which were further used to construct the LR model.

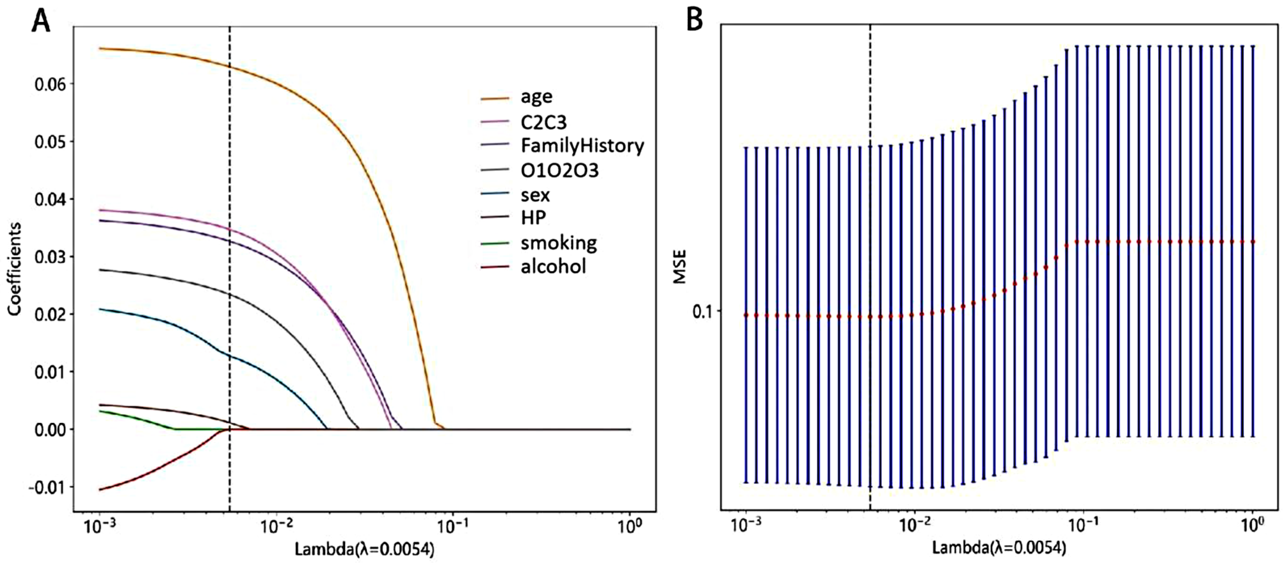

**Fig. 1** Variables selection based on LASSO algorithm for the model. **a** The coefficient profile plot. **b** The cross-validation plot

### Model establishment

Based on the selected features, various machine learning classifiers were used to develop the GC risk prediction models, including LR, NaiveBayes, SVM, KNN, Decision Tree, Random Forests, ExtraTrees, XGBoost, LightGBM, GradientBoosting, and AdaBoost. Most machine learning algorithms showed good diagnostic performance, with AUC values ranging from 0.762 to 0.974 in the training set and 0.608 to 0.812 in the validation set. The specific values for all models are in Table S1. The LR model had the highest AUC value (0.812, 95%CI: 0.737–0.887) in the validation set. In addition, the sample prediction histogram of the LR model showed good calibration, and the DCA curve demonstrated good clinical applicability of the model. The performance of the LR model is in Fig. 2, and other models' AUC values are in Fig. 3. Finally, we constructed a nomogram for GC based on LR (Fig. 4).

## Discussion

To date, the mortality rate of GC remains high in China, which is mainly due to delayed diagnosis. Therefore, it is imperative to improve the early detection of GC [19].

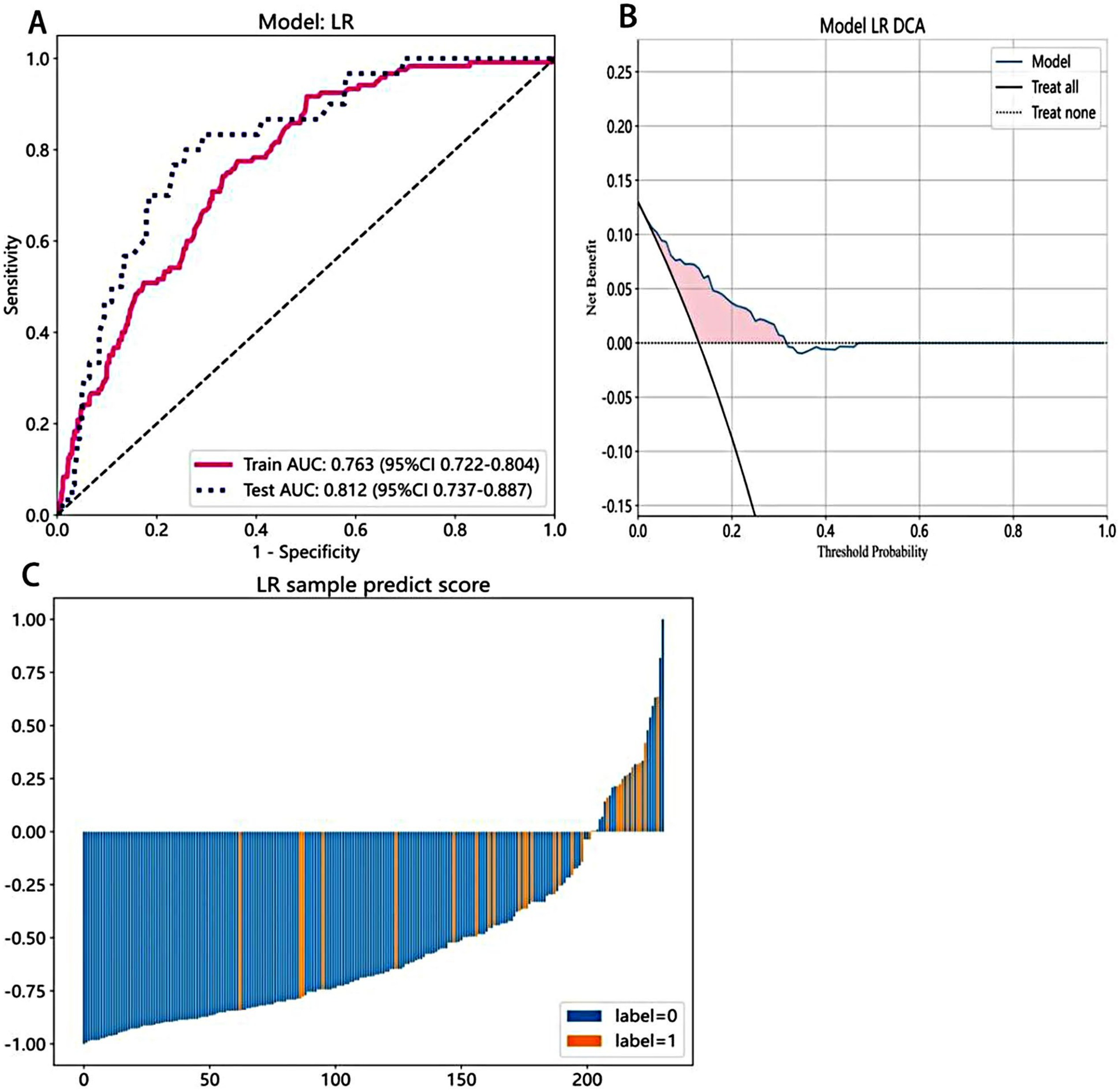

**Fig. 2** The efficacy of the logistic regression (LR) model in the validation set. **a** The receiver operator characteristic (ROC) curves. **b** The decision curve analysis (DCA). **c** The prediction probability histogram

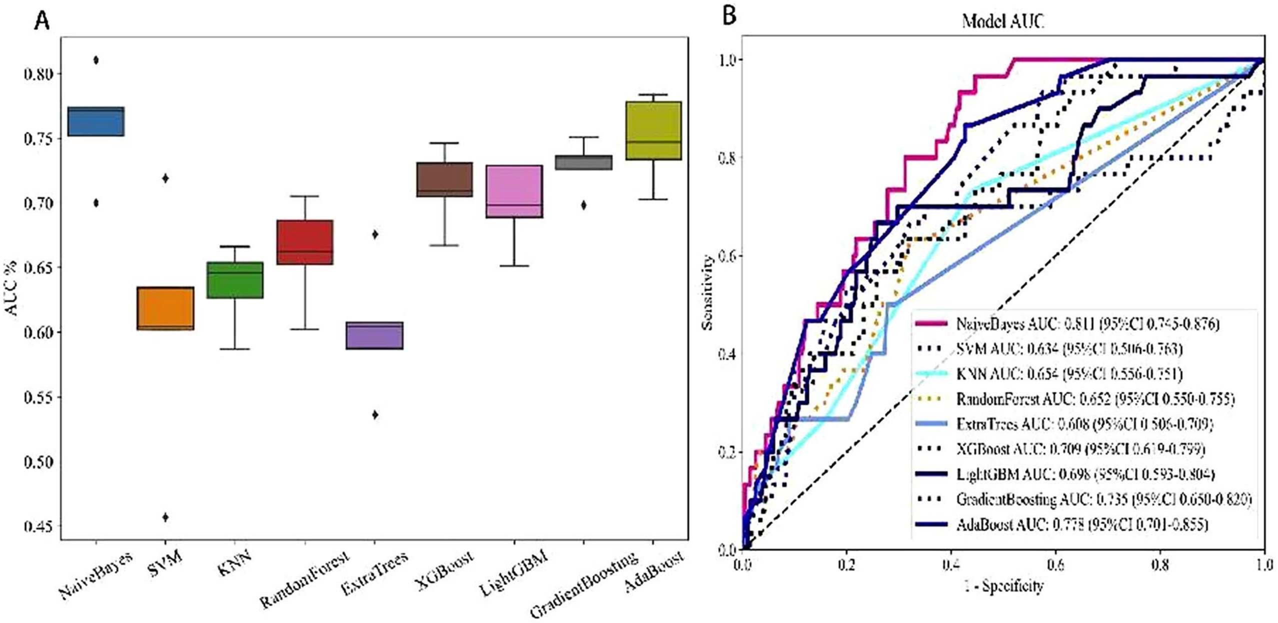

**Fig. 3** The efficacy of other machine learning-based models. **a** The box plot of Area Under the Curve (AUC) and 95%CI in the training set. **b** The Receiver Operator Characteristic (ROC) curves in the validation set

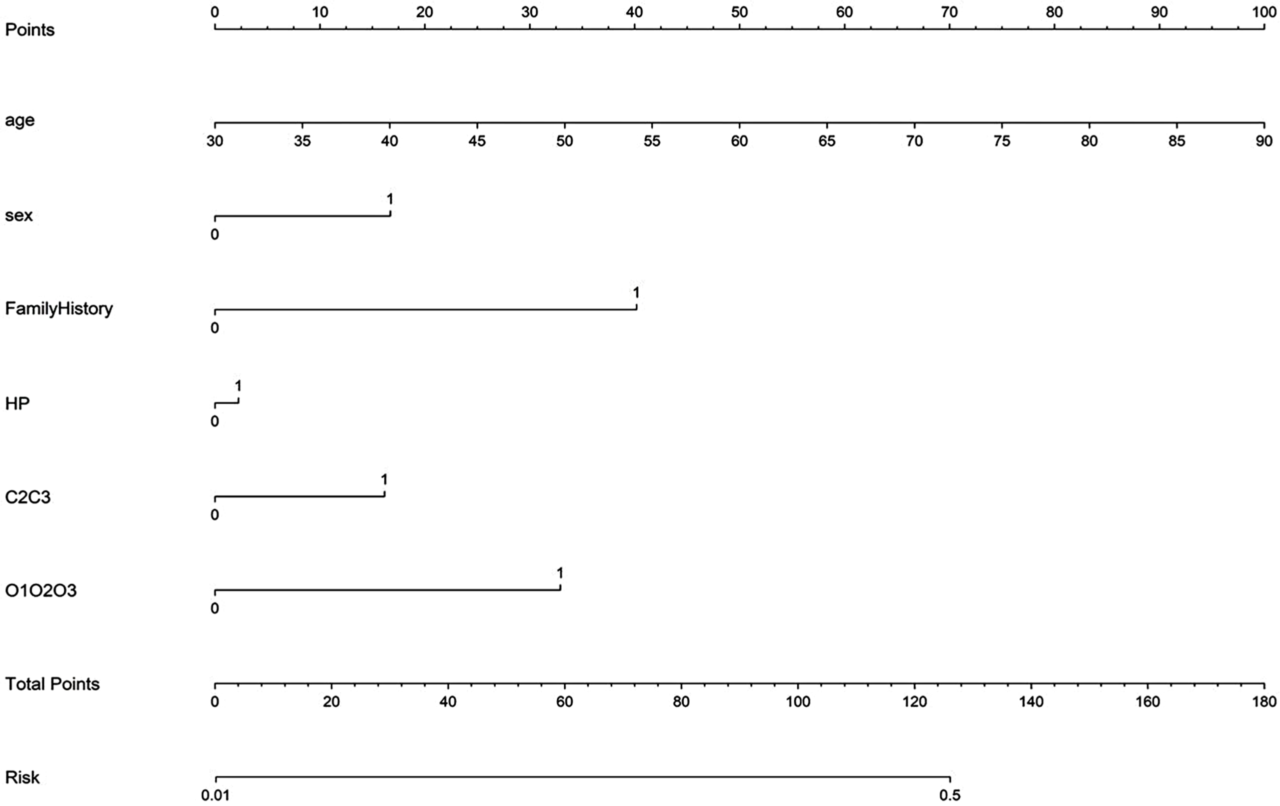

**Fig. 4** The nomogram based on the logistic regression

Currently, the most widely accepted GC risk prediction model in the Chinese population is Lee's Scale [20], which predicts the risk of GC mainly based on serological indicators in asymptomatic individuals. However, there is no GC risk prediction model for Chinese populations using endoscopic atrophy classification.

We believe that the advantage of GC risk assessment under the endoscopy is enabling clinical decisions for high-risk patients, such as determining follow-up time if there were no significant findings during endoscopic examination. In addition, compared to biopsies, it can provide a quantitative assessment of the entire gastric mucosa and improve the accuracy of GC risk assessment.

Much evidence indicated that the incidence of GC increased with age, especially after 40 years old [21, 22]. Compared to females, males have a higher incidence of GC [23, 24]. Smoking [25] and alcohol consumption might also increase the risk of GC. The risk of GC is significantly higher for individuals with a family history of GC. In addition, infection with HP is considered the most significant risk factor for GC. Since the discovery of HP, numerous studies have linked it to GC and precancerous conditions [26, 27]. Therefore, we chose age, sex, family history of GC in first-degree relatives, smoking, alcohol, and HP infection status as predictors, which may have correlations with the occurrence of GC.

In addition, we choose the KTC to assess the degree of gastric atrophy and predict the risk of GC. Although KTC is subjective and may have inter-observer variability, this study simplified KTC into three categories: C0-C1, C2-C3, and O1-O3. The boundary of C1 and C2 is incisura angularis, while cardia for C3 and O1, both with clear markers. This three-category could reduce inter-observer discrepancy and improve the accuracy and reliability of the classification.

LASSO analysis compressed the regression coefficients in the regression equation by generating a penalty function, which could avoid overfitting the model. Therefore, this study adopts LASSO to select variables. After LASSO analysis, we selected five variables for the construction of the model: age, sex, family history of GC, HP infection status, and KTC.

The results of this study showed that there was no significant relationship between smoking or alcohol and GC, which may be because we did not consider the amount of smoking and alcohol consumption. A cohort study [28] conducted on the Singaporean Chinese population indicated that only smoking more than 20 packs per year would increase the risk of GC. We only collected whether the participants smoked or not, without the specific consumption of cigarettes and alcohol consumed, which resulted in the unrelated findings of smoking and alcohol with GC. In the future, more detailed data needs to be collected to verify the relationship of smoking and alcohol with GC in a larger population. Additionally, the nomogram showed a relatively weak importance of HP in the occurrence of GC, which might be associated with precancerous conditions. The incidence of HP-related precancerous diseases, including dysplasia and chronic atrophic gastritis, was a little high in the non-GC group in our study, which reduced the discrepancy of HP between the GC and non-GC group.

There are some advantages in this study. Firstly, we compared the performances of multiple machine learning algorithms to establish models, and the LR model showed the best performance. Subsequently, we developed a nomogram according to the LR equation to display the model, which could assist physicians in assessing the risk of GC based on the endoscopic atrophy classification. However, there were also some limitations in this study. Firstly, we collected the data from a single center without external validation, so we should further validate the model to confirm its generalization. Secondly, we only selected the patients who underwent endoscopy in the study, which may cause selective bias and limit its clinical use. Thirdly, the study was retrospective, and the collected variables were limited. In the future, other endoscopic findings, such as the regular arrangement of the collecting veins, diffuse redness [29], and endoscopic grading of the gastric intestinal metaplasia [30], could be combined to improve the model's performance.

## Conclusion

We established a GC risk prediction model based on endoscopic atrophy classification by multiple machine learning, which can predict the risk of GC and provide guidance for surveillance. However, more populations are needed to validate the model in the future.

### Abbreviations

| | |
|---|---|
| GC | Gastric cancer |
| LASSO | Least absolute shrinkage and selection operator |
| ROC | Receiver Operating Characteristic curve |
| DCA | Decision Curve Analysis |
| KTC | Kimura-Takemoto Classification |
| AUC | Area Under the Curve |
| OLGA | Operative Link on Gastritis Assessment |
| H. pylori | Helicobacter pylori |
| SD | Standard Deviation |
| LR | Logistic Regression |
| SVM | Support Vector Machine |
| KNN | K-Nearest Neighbor |
| AdaBoost | Adaptive Boosting |
| RF | Random Forest |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13860-3.

Supplementary Material 1

**Acknowledgements**
Not applicable.

**Author contributions**
HWX, WJS, and YDL designed the research. YDL, TYZ, BH, SZ, ZYL, YNX and YJS collected the data. MYL, LS and YDL analyzed the data. YDL and QQX wrote the manuscript. HWX, YDL and QQX revised the final manuscript. All authors read and approved the final manuscript.

**Funding**
The study was supported by the National Natural Science Foundation of China, No. 82170650.

**Data availability**
The datasets used during the current study are available from the corresponding author on reasonable request.

## Declarations

**Ethical approval**
This study has been approved by the Ethics Committee of Shandong Provincial Hospital (SWYX: no.2022 – 326). This was a retrospective study and informed consent from patients was waived by the Ethics Committee of Shandong Provincial Hospital.

**Consent for publication**
Not applicable.

**Competing interests**
The authors declare no competing interests.

Received: 11 March 2024 / Accepted: 4 March 2025
Published online: 21 March 2025

### References

1. Sung H, Ferlay J, Siegel RL et al. Global Cancer Statistics. 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians* 2021; 71(3): 209– 49.
2. Salati M, Di Emidio K, Tarantino V, Cascinu S. Second-line treatments: moving towards an opportunity to improve survival in advanced gastric cancer? ESMO Open. 2017;2(3):e000206.
3. Suzuki H, Oda I, Abe S, et al. High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection. Gastric Cancer: Official J Int Gastric Cancer Association Japanese Gastric Cancer Association. 2016;19(1):198–205.
4. Pimentel-Nunes P, Libânio D, Marcos-Pinto R, et al. Management of epithelial precancerous conditions and lesions in the stomach (MAPS II): European society of Gastrointestinal endoscopy (ESGE), European Helicobacter and microbiota study group (EHMSG), European society of pathology (ESP), and sociedade Portuguesa de endoscopia digestiva (SPED) guideline update 2019. Endoscopy. 2019;51(4):365–88.
5. Correa P. Human gastric carcinogenesis: A multistep and multifactorial Process—First American Cancer society award lecture on Cancer epidemiology and Prevention1. Cancer Res. 1992;52(24):6735–40.
6. 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.
7. Dixon MF, Genta RM, Yardley JH, Correa P. Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. *The American journal of surgical pathology* 1996; 20(10): 1161-81.
8. Rugge M, Meggio A, Pennelli G, et al. Gastritis staging in clinical practice: the OLGA staging system. Gut. 2007;56(5):631–6.
9. Hu Y, Bao H, Jin H, et al. Performance evaluation of four prediction models for risk stratification in gastric cancer screening among a high-risk population in China. Gastric Cancer: Official J Int Gastric Cancer Association Japanese Gastric Cancer Association. 2021;24(6):1194–202.
10. Kimura K, Takemoto TJE. An endoscopic recognition of the atrophic border and its significance in chronic gastritis. 1969; 1(03): 87–97.
11. Kono S, Gotoda T, Yoshida S, et al. Can endoscopic atrophy predict histological atrophy? Historical study in united Kingdom and Japan. World J Gastroenterol. 2015;21(46):13113–23.
12. Quach DT, Le HM, Nguyen OT, Nguyen TS, Uemura N. The severity of endoscopic gastric atrophy could help to predict operative link on gastritis assessment gastritis stage. J Gastroenterol Hepatol. 2011;26(2):281–5.
13. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nat Rev Mol Cell Biol. 2022;23(1):40–55.
14. Myszczynska MA, Ojamies PN, Lacoste AMB, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat Reviews Neurol. 2020;16(8):440–56.
15. Liu WZ, Xie Y, Lu H, et al. Fifth Chinese National consensus report on the management of Helicobacter pylori infection. Helicobacter. 2018;23(2):e12475.
16. Kimura K, Takemoto TJE. An Endoscopic Recognition of the Atrophic Border and its Significance in Chronic Gastritis. 1969; 1: 87–97.
17. Schlemper RJ, Riddell RH, Kato Y, et al. The Vienna classification of Gastrointestinal epithelial neoplasia. Gut. 2000;47(2):251–5.
18. Dixon MF. Gastrointestinal epithelial neoplasia: Vienna revisited. Gut. 2002;51(1):130–1.
19. Zong L, Abe M, Seto Y, Ji J. The challenge of screening for early gastric cancer in China. Lancet (London England). 2016;388(10060):2606.
20. Cai Q, Zhu C, Yuan Y, et al. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population: a nationwide multicentre study. Gut. 2019;68(9):1576–87.
21. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer epidemiology, biomarkers & prevention: a publication of the American association for Cancer research*. Cosponsored Am Soc Prev Oncol. 2014;23(5):700–13.
22. Yao K, Uedo N, Kamada T, et al. Guidelines for endoscopic diagnosis of early gastric cancer. Dig Endoscopy: Official J Japan Gastroenterological Endoscopy Soc. 2020;32(5):663–98.
23. González CA, Agudo A. Carcinogenesis, prevention and early detection of gastric cancer: where we are and where we should go. Int J Cancer. 2012;130(4):745–53.
24. Leung WK, Wu MS, Kakugawa Y, et al. Screening for gastric cancer in Asia: current evidence and practice. Lancet Oncol. 2008;9(3):279–87.
25. Ferro A, Morais S, Rota M, et al. Tobacco smoking and gastric cancer: meta-analyses of published data versus pooled analyses of individual participant data (StoP Project). Eur J cancer Prevention: Official J Eur Cancer Prev Organisation (ECP). 2018;27(3):197–204.
26. Nie Y, Wu K, Yu J, et al. A global burden of gastric cancer: the major impact of China. Expert Rev Gastroenterol Hepatol. 2017;11(7):651–61.
27. Shakir SM, Shakir FA, Couturier MR. Updates to the diagnosis and clinical management of Helicobacter pylori infections. Clin Chem. 2023;69(8):869–80.
28. Lee JWJ, Zhu F, Srivastava S, et al. Severity of gastric intestinal metaplasia predicts the risk of gastric cancer: a prospective multicentre cohort study (GCEP). Gut. 2022;71(5):854–63.
29. Sugimoto M, Ban H, Ichikawa H, et al. Efficacy of the Kyoto classification of gastritis in identifying patients at high risk for gastric Cancer. Intern Med (Tokyo Japan). 2017;56(6):579–86.
30. Pimentel-Nunes P, Libânio D, Lage J, et al. A multicenter prospective study of the real-time use of narrow-band imaging in the diagnosis of premalignant gastric conditions and lesions. Endoscopy. 2016;48(8):723–30.

## Publisher's note