RESEARCH ARTICLE

Discrimination between Precancerous Gastric Lesions and Gastritis Using a Gastric Cancer Risk Stratification Model

John D Murphy[1]*, Meira Epplein[2], Feng-Chang Lin[3], Melissa A Troester[1], Hazel B Nichols[1], Julia Butt[4], Kaifeng Pan[5], Weicheng You[5], Andrew Olshan[1]

Abstract

Background: Seropositivity to certain Helicobacter pylori proteins may affect development of gastric lesions that could become cancerous. Previously, we developed a model of gastric cancer risk including gender, age, HP0305 sero-positivity, HP1564 sero-positivity, UreA antibody titer and serologically defined chronic atrophic gastritis (termed: "Lasso model"). **Methods:** We evaluated the Lasso model's ability to discriminate individuals with precancerous gastric lesions (n=320) from individuals with superficial or mild atrophic gastritis (n=226) in Linqu County, China, a population at high risk for gastric cancer. We also compared its performance to the ABC Method, a gastric cancer risk stratification tool currently used in East Asia. **Results:** For distinguishing precancerous lesions from those with gastritis, the receiver operating characteristic curve had an area under the curve (AUC) of 73.41% (95% CI: 69.10%, 77.71%) and, at Youden's Index, a sensitivity of 78.44% (59.38%, 82.50%) and specificity of 64.72% (95% CI: 58.85%, 81.42%). Positive predictive value (PPV) was 75.38% (72.78%, 82.51%). Specificity, AUC and PPV were significantly greater (p < 0.05) than those of the ABC Method. When specificity was held constant, the Lasso model had greater sensitivity, PPV and negative predictive value (NPV) than the ABC Method. However, adjusting the ABC Method for age and gender negated the Lasso model's significant improvement in AUC. **Conclusions:** The Lasso model for gastric cancer risk prediction can classify precancerous lesions with significantly greater AUC than the ABC Method and, at constant specificity, with greater sensitivity, PPV and NPV. However, adding age and gender to the ABC Method, as included in the Lasso model, substantially improved its performance and negated the Lasso model's advantage.

Keywords: Intestinal metaplasia- dysplasia- early detection- machine learning- gastric cancer- predictive modelling

Introduction

About one million incident cases of, and over 750,000 deaths from, gastric cancer occur every year (Ferlay et al., 2015; Sung et al., 2021). The disease is especially common in East Asia, including China, Japan, and Korea (Ferlay et al., 2015; Sung et al., 2021). Gastric cancer's high mortality arises in large part from its frequently being diagnosed at late stage when treatment is difficult (Allemani et al., 2015; Colquhoun et al., 2015). Therefore, improvements in early detection, and interception of pre-malignant lesions, may reduce mortality and morbidity from this disease.

To non-invasively identify who would benefit from screening, Yamaguchi et al., (2016) created the ABC Method for gastric cancer risk stratification in Japan. The ABC Method has two serological components: infection with the bacterium *Helicobacter pylori* (*H.pylori*) and pepsinogen-defined chronic atrophic gastritis (CAG). However, the ABC Method does not account for differences in risk conferred by different *H. pylori* virulence factors (Sasazuki et al., 2006). In previous work, we used the Least Absolute Shrinkage and Selection Operator (Lasso) to develop a predictive model that incorporated antibody response to three *H. pylori* proteins (HP 0305, HP 1564 and UreA), pepsinogen-defined CAG, age and gender (Murphy et al., 2022; Tibshirani, 1996). This model was built using data from three cohorts of the *H. pylori* Biomarker Cohort Consortium (HpBCC) that collected serological *H. pylori* and pepsinogen data: the Japan Public Health Center Study (JPHC) I and JPHC II in Japan and the Linxian Nutrition Intervention Trial (NIT) in China (Cai et al., 2016). The model achieved an area under the receiver-operating characteristic curve of 73.37% (95% CI: 70.42%, 76.32%) and a sensitivity of 73.56% (95% CI: 69.86%, 77.27%) (Murphy et al.,

*[1]University of North Carolina at Chapel Hill, Gillings School of Global Public Health, Department of Epidemiology, Chapel Hill, NC, USA. [2]Duke University, Department of Population Health Sciences, and Duke Cancer Institute, Cancer Risk, Detection, and Interception Program, Durham, NC, USA. [3]University of North Carolina at Chapel Hill, Gillings School of Global Public Health, Department of Biostatistics, Chapel Hill, NC, USA. [4]Deutsches Krebsforschungszentrum, Heidelberg, Germany. [5]Peking University Cancer Hospital, Beijing, China. *For Correspondence: jack.murphy@nih.gov*

2022). However, the model was not assessed for ability to distinguish precancerous lesions from gastritis.

Precancerous gastric lesions, including metaplasia and dysplasia, are steps on a cascade of gastric carcinogenesis and can be treated by resection and/or *H. pylori* eradication (Correa et al., 2007; Correa et al., 2012; Correa et al., 2010). A non-invasive test that accurately distinguished such lesions from milder stomach conditions such as gastritis could improve their interception and thus possibly reduce gastric cancer incidence. In addition, individuals with precancerous lesions could be targeted for ablation or *H. pylori* eradication therapy, the latter of which has been shown to reverse the growth of precancerous lesions and reduce gastric cancer risk (Doorakkers et al., 2016; Sugano, 2019). To discern precancerous lesions from superficial gastritis, Epplein et al., (2018) constructed a model in the Linqu County Study incorporating age, gender, smoking status, *H. pylori* infection status, and seropositivity to two *H. pylori* proteins: HP 0305 and HP 1564. This model had an area under the receiver operator characteristic curve of 75.10% (95% CI: 72.45%, 77.74%). Adding pepsinogen to this model in a sensitivity analysis did not significantly or substantially change the AUC (Epplein et al., 2018).

In the present study, we assessed the ability of the Lasso-derived gastric cancer risk model developed from a consortium of prospective cohort studies in East Asia (the Helicobacter pylori Biomarker Cohort Consortium, HpBCC), which incorporated pepsinogen levels, to discriminate between cases of precancerous lesions and individuals with superficial gastritis or mild CAG using data from the Linqu County Study (Murphy et al., 2022). Its performance could shed more light on the added value of seropositivity to *H. pylori* proteins (including UreA) and chronic atrophic gastritis status to discriminate precancerous gastric lesions from superficial gastritis.

Materials and Methods

Study Population

The Linqu County Study was originally conceived in 2002 as a double-blind, randomized placebo-controlled trial to assess the efficacy of celecoxib – a COX-2 inhibitor – in either preventing incident gastric cancer or promoting the regression of precancerous gastric lesions (Wong et al., 2012). Out of 3161 Linqu County residents assessed for eligibility, 2813 agreed to participate in the initial screening. The trial included participants aged 35-64 years who were positive for *H. pylori* infection and had a baseline histology of severe CAG, intestinal metaplasia, indefinite dysplasia or dysplasia (Wong et al., 2012). Exclusion criteria before randomization for the intervention trial were: refusal to provide informed consent, a previous negative *H. pylori* test, non-atrophic gastritis, mild or moderate CAG, heart failure, emphysema, liver or renal disease, bleeding diathesis and/or requiring anticoagulant therapy, hypertension (defined as diastolic blood pressure >95 mm Hg, systolic blood pressure >165 mm Hg), history of stroke or transient ischemic attack within the last two years, history of neoplastic disease within the previous ten years, or allergy to antibiotics (Wong et al., 2012). For this study, 1402 participants who underwent upper endoscopy at baseline were considered for analysis. Among these, 546 individuals had valid pepsinogen measurements and therefore comprised the final analysis set.

Outcome

In this analysis, we applied a predictive model for gastric cancer risk to detect precancerous gastric lesions. This was a composite endpoint, where cases of precancerous lesions were defined as individuals who had one of the following three diagnoses (n=320): intestinal metaplasia, indefinite dysplasia, or dysplasia. Individuals with gastric cancer were not included in this analysis. There is evidence that precancerous gastric lesions are markers along a path of gastric carcinogenesis known as the "Correa Cascade," which extends from normal gastric mucosa to chronic gastritis, then chronic atrophic gastritis, intestinal metaplasia, dysplasia and, finally, adenocarcinoma (Correa et al., 2007; Correa et al., 2012; Correa et al., 2010). In the Linqu County Study, very few participants were found to have normal gastric mucosa (Epplein et al., 2018; Wong et al., 2012). Therefore, a random sample of participants diagnosed with superficial gastritis or mild CAG at baseline were chosen to comprise the referent group (n=226). Individuals with superficial gastritis or mild CAG could be at higher risk of developing precancerous lesions than individuals with normal gastric mucosa, which may give a more valid estimate of the test's specificity.

Outcome status was determined by upper endoscopy examination. This was performed by four experienced gastroenterologists using fiber-optic or video endoscopes (Wong et al., 2012). The gastroenterologists examined the gastric mucosa and took five biopsy samples from standard sites according to the Updated Sydney System (Dixon et al., 1996): lesser curvature of the antrum, greater curvature of the antrum, angulus, lesser curvature of the body, and greater curvature of the body (Wong et al., 2012). Biopsy specimens were diagnosed separately, then each participant was assigned a global diagnosis based on the most severe diagnosis from among their respective specimens. Three senior pathologists in the Department of Pathology at the Peking University School of Oncology reviewed each slide to make a diagnosis according to the Updated Sydney System and Padova International Classification (Dixon et al., 1996; Rugge et al., 2000). 196 slides were blindly tested for quality control by a pathologist at a separate institution. Consensus was reached for 188 (95.9%) of the slides (Wong et al., 2012).

Features of the Predictive Model
H. pylori multiplex serology

Antibody responses to *H. pylori* antigens were assessed by using multiplex serology following the method developed by Michel et al., (2009). The multiplex assay quantifies antibody responses to recombinantly expressed glutathione-S-transferase (GST) fusion proteins via an immunosorbent assay combined with fluorescent bead technology (Luminex) (Epplein et al., 2018; Michel et al., 2009). This combined assay can detect human IgA, IgM, and IgG antibodies to 13 *H. pylori* proteins (UreA,

Catalase, GroEL, NapA, CagA, HP0231, VacA, HpaA, Cad, HyuA, HP1564, HcpC and HP0305). Antigen-specific thresholds were calculated as the mean of the median MFI plus three times the standard deviation (excluding positive outliers) using 17 *H. pylori*-negative sera. In the derivation data set, 13 *H. pylori* variables listed above were assessed in both linear MFI value (among *H. pylori*-seropositive participants only) and binary functional form for potential inclusion as features of the new predictive model (Murphy et al., 2022).

Assessment of Serum Pepsinogen Levels

Pepsinogen levels were assessed by ELISA kits (Eagle Biosciences) (Epplein et al., 2018).In brief, 25 µL of serum for pepsinogen I and 50 µL of serum for pepsinogen II were applied in duplicate to a microplate coated in streptavidin. Next, the samples were incubated with their respective biotinylated capture and horseradish peroxidase (HRP)-labelled tracer antibodies and HRP substrate was added for signal detection. Stop solution was used to stop the reaction and the absorbance was measured in a microplate reader at 450 nm. Assay standards were then run on each plate to construct a plate-specific standard curve for determination of each sample's concentration (ng/mL) of pepsinogens I and II. To ensure reliability of the assay, two control samples with known pepsinogen I and II concentrations provided by the manufacturer were applied to each plate. Linear pepsinogen I and II variables, as well as the ratio of pepsinogen I:II (linear), were evaluated as potential features of the predictive model (Murphy et al., 2022).In addition, a binary variable for serologically defined chronic atrophic gastritis (CAG) was evaluated with the following coding: CAG-positive for individuals with pepsinogen I concentration ≤ 70 µg/L and pepsinogen I:II ratio ≤ 3.0, CAG-negative otherwise.

Model Building

Previously, in a subset of the HpBCC (n=1402), as mentioned above, we used the least absolute shrinkage and selection operator (Lasso) to build a predictive model for gastric cancer (Murphy et al., 2022; Tibshirani, 1996). Lasso is a penalized regression technique that selects variables by fitting a risk model with the constraint that the sum of the absolute values of the regression coefficients cannot exceed a pre-determined threshold (Liu et al., 2019; Tibshirani, 1996; Tibshirani, 1997). This threshold excludes variables that contribute least to predicting the outcome. The remaining covariates are thus selected as predictors of gastric cancer. We decided to use the same features in this study because gastric adenocarcinoma tends to follow a progressive cascade from *H. pylori* infection, through metaplasia, dysplasia, then finally cancer (Sutton et al., 2010).Therefore, it could be that seropositivity to the *H. pylori* proteins associated with gastric cancer are also relevant to precursor lesion development.

The ABC Method was constructed as a categorical variable with the following three levels: A (referent; *H. pylori*-negative, CAG-negative), B (*H. pylori*-positive, CAG-negative) and C+D (*H. pylori*-positive, CAG-positive OR *H. pylori*-negative, CAG-positive). The ABC

Method was coded identically in the present application study. *H. pylori* seropositivity for the ABC Method was also defined as seropositivity to ≥4 *H. pylori* proteins. In addition, we adjusted the ABC Method model for gender and age in order to make it more comparable to our Lasso-derived model. Older age and male gender are both associated with higher gastric cancer risk, which may mean they are also associated with the development of precancerous gastric lesions (Sung et al., 2021).

Statistical Analysis

In the derivation set, the association between incident gastric cancer and the features selected by Lasso was estimated using Cox proportional hazards regression (Tibshirani, 1996; Tibshirani 1997).The Linqu County Study data set, however, is cross-sectional. Therefore, we used logistic regression to generate odds ratios for the association between precancerous lesions and each feature of the Lasso-generated model in the present analysis.

The model's predictive accuracy was estimated using receiver-operator characteristic (ROC) curves. Area under the curve (AUC), sensitivity, specificity, positive predictive value, and negative predictive value are reported.

All statistical analyses were carried out in R statistical software (version 4.0.2) using RStudio and the glmnet and pROC packages (Friedman et al., 2010; Robin et al., 2011; RStudio, 2020; Team, 2016).

Results

The final model selected by Lasso incorporated the following six variables: gender (binary); age (linear; centered at 57 years, the median age for the data set), UreA (linear), HP 0305 (binary), HP 1564 (binary), and serologically defined CAG (binary).

There were 320 cases of precancerous lesions and 226 referent individuals in our analysis set (Table 1). Cases were more likely than referent individuals to be men (53% vs. 42%), be *H. pylori* sero-positive (83% vs. 54%), have chronic atrophic gastritis (7% vs. 3%) and to be current smokers (47% vs. 35%). Cases were generally older than referent participants and had a lower median pepsinogen I/II ratio (7.6 vs. 14.0). There was no difference in family history of gastric cancer by precancerous lesion status (Table 1).

Of the six predictors in the Lasso model developed for gastric cancer risk prediction, four were also significantly associated with odds of gastric precancerous lesions. The strongest association was observed between HP 1564 and precancerous gastric lesions (OR=3.77; 95% CI: 2.27, 6.33; Table 2). Overall, the Lasso model's AUC in this application study was very similar to its AUC within the derivation set (73.79%; 95% CI: 70.86%, 76.73%). In addition, its sensitivity was higher than that achieved in the derivation set (73.56%; 95% CI: 69.86%, 77.27%), although with lower precision. For the ABC Method, individuals in Groups B or C+D had a 4- to 7-fold increase in odds of precancerous lesions than individuals in Group A (Table 2).

We compared the Lasso model's ability to discriminate

Table 1. Demographic Characteristics of Application Population, Linqu County, Shandong Province, China, 2002–2004 (N = 546).

	Referent†	Case: Metaplasia	Case: Indefinite Dysplasia	Case: Dysplasia	Total Cases of precancerous lesions
n	226	142	116	62	320
Gender (%)					
Women	130 (57.5)	79 (55.6)	43 (37.1)	28 (45.2)	150 (46.9)
Men	96 (42.5)	63 (44.4)	73 (62.9)	34 (54.8)	170 (53.1)
H. pylori status (%)					
H. pylori–	103 (45.6)	20 (14.1)	23 (19.8)	11 (17.7)	54 (16.9)
H. pylori+	123 (54.4)	122 (85.9)	93 (80.2)	51 (82.3)	266 (83.1)
Age category, years (%)					
≤40	28 (12.4)	9 (6.3)	9 (7.8)	2 (3.2)	20 (6.2)
41-50	106 (46.9)	67 (47.2)	53 (45.7)	31 (50.0)	151 (47.2)
51-60	75 (33.2)	56 (39.4)	41 (35.3)	22 (35.5)	119 (37.2)
>60	17 (7.5)	10 (7.0)	13 (11.2)	7 (11.3)	30 (9.4)
Serum Pepsinogen I (μg/L) Median (IQR)	32.50 (23.00, 50.00)	33.50 (25.00, 56.00)	36.50 (27.00, 50.00]	36.00 (25.25, 61.75)	35.00 [26.00, 53.25]
Median Pepsinogen I/II Ratio Median (IQR)	14.00 (8.08, 21.00)	7.00 (4.75, 11.11)	7.50 (4.88, 12.45)	9.52 (5.88, 16.00)	7.55 [5.00, 12.41]
Chronic Atrophic Gastritis‡ (%)					
CAG–	220 (97.3)	134 (94.4)	104 (89.7)	61 (98.4)	299 (93.4)
CAG+	6 (2.7)	8 (5.6)	12 (10.3)	1 (1.6)	21 (6.6)
Smoking status (%)					
Not Current Smoker	148 (65.5)	92 (64.8)	47 (40.5)	31 (50.0)	170 (53.1)
Current Smoker	78 (34.5)	50 (35.2)	69 (59.5)	31 (50.0)	150 (46.9)
Family History (%)					
No Family History of Gastric Cancer	217 (96.0)	139 (97.9)	110 (94.8)	58 (93.5)	307 (95.9)
Family History of Gastric Cancer	9 (4.0)	3 (2.1)	6 (5.2)	4 (6.5)	13 (4.1)

†, Referent group comprises those with superficial gastritis or mild chronic atrophic gastritis; ‡, Defined as pepsinogen I ≤ 70 μg/L and pepsinogen I/II ratio ≤ 3

Table 2. Summary of the Parameter Estimates of the Lasso Model and ABC Method for Classification of Gastric Precancerous Lesions in the Linqu County Study Data Set (N = 546; 226 Referent, 320 Cases)

Parameter[†]	Coefficient (SE)	Odds Ratio (95% CI)
Lasso Model		
Gender (men; binary)	0.468 (0.177)	1.60 (1.10, 2.33)
Age[‡] (linear)	0.039 (0.015)	1.04 (1.01, 1.07)
UreA[§] (linear)	−0.00002 (0.00007)	1.00 (1.00, 1.00)
HP 0305 (seropositive; binary)	0.422 (0.255)	1.43 (0.92, 2.51)
HP 1564 (seropositive; binary)	1.330 (0.261)	3.77 (2.27, 6.33)
CAG[¶] (seropositive; binary)	0.734 (0.508)	2.08 (0.81, 6.13)
ABC Method[#]		
A (*H. pylori*–, CAG–)	REF	REF
B (*H. pylori*+, CAG–)	1.46 (0.206)	4.30 (2.89, 6.48)
C + D (*H. pylori*+, CAG+; or *H. pylori*–, CAG+)	1.97 (0.494)	7.14 (2.86, 20.50)

[†], All Lasso parameters are adjusted for each other; [‡], Centered at 49 years (median age in the data set); [§], Median reporter fluorescence intensity values < 74 were considered sero-negative and recoded as 0; [¶], Defined as pepsinogen I ≤ 70 μg/L and pepsinogen I/II ratio ≤ 3; [#], Separate regression equation from Lasso model parameters.

between cases of precancerous lesions and referent participants by plotting ROC curves and reporting the area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV; Table 3). Youden's Index was used to determine the threshold for sensitivity and specificity. The Lasso model's AUC (73.42%; 95% CI: 69.12%, 77.71%) was significantly (DeLong's p < 0.001) greater than that of the ABC Method 65.43% (95% CI: 61.50%, 69.37%). Additionally, the Lasso model had higher specificity (63.72%; 95% CI: 58.41%, 82.30%) and PPV (75.38%;

95% CI: 72.61%, 82.49%) than the ABC Method (Table 3). The ABC Method achieved a higher sensitivity (84.38%; 95% CI: 80.31%, 88.12%) than the Lasso model (78.44%; 95% CI: 59.38%, 82.50%). The NPVs of both models were very similar (Table 3).

At 75% specificity, the Lasso model compared to the ABC Method had greater sensitivity (78.44% vs. 39.86%), PPV (75.38% vs. 69.30%) and NPV (67.61% vs. 46.83%) (Table 4). The same trends were observed at 85% specificity (Lasso vs. ABC Method, sensitivity: 45.59% vs. 24.96%; PPV: 81.15%, 70.20%; NPV: 52.46%
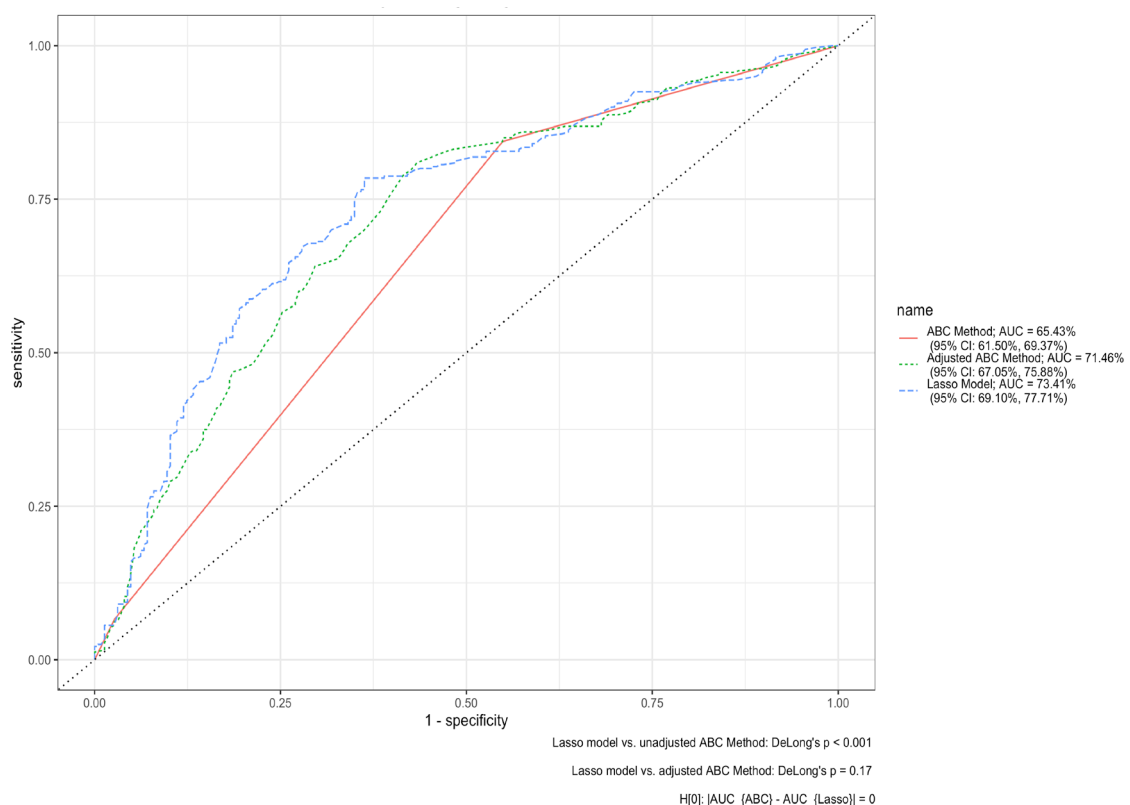

Figure 1. Receiver-Operating Characteristic Curves Showing the Unadjusted ABC Method, adjusted ABC Method (for Gender and Age), and Lasso Model's discrimination capability for precancerous gastric lesions vs. superficial gastritis or mild atrophic gastritis.

Table 3. Lasso Model and ABC Method Classification Capability for Precancerous Gastric Lesions (n = 546)

Model Name	AUC[†] (%)	p[#]	Sensitivity[††] (%)	Specificity[††] (%)	PPV[††] (%)	NPV[††] (%)
ABC Method[‡]	65.43 (61.50, 69.37)	< 0.001	84.38 (80.31, 88.12)	45.13 (38.94, 51.77)	68.53 (65.85, 71.35)	67.12 (60.76, 73.53)
Adjusted ABC Method[§]	71.46 (67.05, 75.88)	0.17	82.81 (78.75, 86.88)	52.65 (46.02, 59.29)	71.24 (68.30, 74.37)	68.39 (62.70, 74.32)
Lasso Model[¶]	73.42 (69.12, 77.71)	N/A	78.44 (58.44, 82.50)	63.72 (58.41, 82.30)	75.38 (72.61, 82.94)	67.61 (57.76, 72.59)

[†], Area under the receiver operating characteristic curve (95% Confidence Intervals from DeLong's Method); [‡], Three levels: A (*H. pylori*–, Chronic atrophic gastritis (CAG)–), B (*H. pylori*+, CAG–, C (*H. pylori*+, CAG+; or H. pylori–, CAG+); [§], Three levels: A (*H. pylori*–, Chronic atrophic gastritis (CAG)–), B (*H. pylori*+, CAG–, C (*H. pylori*+, CAG+; or H. pylori–, CAG+). Adjusted for age (linear, in years) and gender (binary); [¶], Six predictors: gender (binary), age (continuous; centered at 57 years), UreA (continuous), HP 0305 (binary), HP 1564 (binary), serologically defined CAG (binary); [#], Comparing Lasso model to ABC Method and, separately, Lasso model with Adjusted ABC Method. DeLong's test of the null hypothesis that |AUCLasso – AUCABC| = 0; [††], Threshold of predicted probability of precancerous lesions at 53.79% chosen by Youden's Index = (sensitivity + specificity) –1. 95% Confidence Intervals produced from bootstrap with 10,000 iterations

vs. 44.44%). At 95% specificity, both models performed very similarly.

However, when age and gender were added to the ABC Method, its AUC improved substantially to 71.46% (95% CI: 67.05, 75.88%). The difference in AUC between the Lasso model and ABC Method thus became non-significant (DeLong's p = 0.17). The Lasso model had higher specificity (63.72%; 95% CI: 58.41%, 82.30%) and PPV (75.38%; 95% CI: 72.61%, 82.49%) than the ABC Method (Table 3). The ABC Method, adjusted for age and gender, achieved a higher sensitivity (82.81%; 95% CI: 78.75%, 86.88%) than the Lasso model (78.44%; 95% CI: 59.38%, 82.50%). The NPVs of both models were very similar (Table 3). Figure 1 shows the receiver operating characteristic curve comparing the Lasso model to the ABC Method and, separately, the ABC Method plus age and gender.

At 75%, 85%, and 95% specificity, the Lasso model also exhibited higher sensitivity and NPV than the ABC Method adjusted for age and gender, but the magnitude of the difference was smaller than that between the Lasso model and the unadjusted ABC Method (Table 4).

Sensitivity Analysis

Only three features of the Lasso-derived model (age, gender and HP 1564) were significantly positively associated with precursor lesions (Table 2). We decided to explore how a classification model containing only these three variables would perform against the ABC Method.

The reduced model had an AUC of 73.38% (95% CI: 69.07%, 77.70%), which was almost identical to the full Lasso model's AUC (Table 3) and borderline greater than the adjusted ABC Method's AUC (DeLong's p = 0.05). At the Youden's Index threshold, the reduced model displayed a sensitivity of 73.13% (95% CI: 57.50%, 77.81%), a specificity of 69.47% (95% CI: 64.16%, 83.63%), a PPV of 77.23% (95% CI: 74.10%, 84.00%) and NPV of 64.61% (95% CI: 57.51%, 69.36%). The reduced model thus showed greater specificity than the full Lasso model but lower sensitivity, PPV and NPV.

Discussion

The Lasso model comprising age, gender, seropositivity to HP 1564, HP 0305, and Urea, and serologically defined chronic atrophic gastritis classified precancerous gastric lesions at a similar level of performance in this application data set (Linqu County) to how it classified gastric cancer within the derivation data set (HpBCC). This suggests the Lasso model could be applied to detect precancerous lesions as well as actual cancer. Additionally, the new model seems to transport to different East Asian populations with low variability. However, the Lasso model did not have a significantly greater AUC than the ABC Method when the latter was adjusted for age and gender. In terms of classifying precancerous lesions, while the Lasso model had a slightly lower sensitivity than the ABC Method, it had substantially higher specificity. The

Table 4. Lasso Model and ABC Method Classification Capability for Precancerous Gastric Lesions at Pre-determined Specificity Levels (n = 546)

Specificity	Model Name	Sensitivity[¶] (%)	PPV[¶] (%)	NPV[¶] (%)
75%	ABC Method[†]	39.86 (34.86, 45.70)	69.30 (66.38, 72.13)	46.83 (44.85, 49.38)
	Adjusted ABC Method[‡]	55.94 (45.12, 66.29)	76.01 (71.87, 78.97)	54.59 (49.11, 61.11)
	Lasso Model[§]	78.44 (58.44, 82.50)	75.38 (72.61, 82.94)	67.61 (57.76, 72.59)
85%	ABC Method[†]	24.96 (20.82, 29.31)	70.20 (66.27, 73.45)	44.44 (43.12, 45.92)
	Adjusted ABC Method[‡]	37.50 (28.03, 49.34)	77.97 (72.57, 82.33)	48.99 (45.48, 54.23)
	Lasso Model[§]	45.59 (34.98, 58.44)	81.15 (76,75, 84.65)	52.46 (48.00, 59.09)
95%	ABC Method[†]	10.06 (5.84, 13.89)	74.01 (62.32, 79,73)	42.73 (41.61, 43.79)
	Adjusted ABC Method[‡]	15.19 (5.22, 26.34)	81.14 (59.64, 88.18)	44.17 (41.45, 47.67)
	Lasso Model[§]	16.13 (0.06, 29.22)	82.03 (62.46, 89.22)	44.44 (41.62, 48.66)

[†], Three levels: A (H. pylori–, Chronic atrophic gastritis (CAG)–), B (H. pylori+, CAG–, C (H. pylori+, CAG+; or H. pylori–, CAG+); [‡], Three levels: A (H. pylori–, Chronic atrophic gastritis (CAG)–), B (H. pylori+, CAG–, C (H. pylori+, CAG+; or H. pylori–, CAG+). Adjusted for age (linear, in years) and gender (binary); [§], Six predictors: gender (binary), age (continuous; centered at 57 years), UreA (continuous), HP 0305 (binary), HP 1564 (binary), serologically defined CAG (binary); [¶], 95% Confidence Intervals produced from bootstrap with 10,000 iterations

considerably higher specificity achieved by the Lasso model in detecting precancerous lesions suggests that the new model may be a more appropriate test than the ABC Method to apply at an early, general population-level stage of screening for precancerous gastric lesions.

When applying a test for classifying precancerous lesions to the general population, very few cases will develop into life-threatening disease, consequently high specificity should generally be prioritized over high sensitivity (Cole et al., 1980). A low specificity test could increase the cost on the healthcare system from overtreatment, overburden diagnostic services, and discourage people from participating in repeat screening (Cole et al., 1980). Therefore, it is promising that the Lasso model's specificity was over 10% greater than the ABC Method's in this study. Repeat screening with the Lasso model may also improve its sensitivity.

An advantage of the Lasso's high specificity, in tandem with the high prevalence of precancerous lesions in Linqu County, was the high PPV observed in this study (Cole et al., 1980). Individuals who test positive under the Lasso model could be targeted for *H. pylori* eradication treatment, which has been shown to reduce the incidence of gastric cancer. Alternatively, individuals with precancerous lesions could undergo ablation or further surveillance. This could ultimately intercept precancers and possibly prevent gastric cancer from developing. However, clinical studies of the Lasso model will be necessary before drawing such conclusions. Additionally, the Lasso model's PPV was only slightly higher than that of the ABC Method at Youden's Index and at high levels of specificity.

Our new model achieved a very similar AUC to another classification model of precancerous lesions developed in the Linqu County Study comprised of age, smoking status (current vs. ever/never), *H. pylori* seropositivity, HP 0305 seropositivity and HP 1564 seropositivity (AUC= 75.10%; 95% CI: 72.45%, 77.74%) (Epplein et al., 2018). Serologically defined CAG was not included in that model, which suggests that, despite being a very strong predictor of gastric cancer in the derivation data set, CAG did not improve classification beyond what could be achieved by HP 0305 and HP 1564 sero-positivity status. This may be because, in the Linqu County Study, the referent group included individuals with mild CAG or superficial gastritis. Moreover, when we explored the discrimination capability of a reduced Lasso model that only contained the statistically significant features age, gender and HP 1564, it performed very similarly to the full Lasso model. This suggests that seropositivity to HP 1564 is an especially strong predictor of gastric precancerous lesions. However, the relatively small sample size of the Linqu County data set meant our estimated odds ratios for each predictor were somewhat imprecise.

In the derivation data set, we also examined the Lasso model's discrimination performance within strata of study site (Murphy et al., 2022). In the Linxian Nutrition Intervention Trial, which was performed in another region of China that carries a high burden of gastric cancer, the Lasso model displayed a similar sensitivity (81.17%; 95% CI: 76.03%, 86.30%; n = 633) to what we observed in this

application study (Li et al., 1993; Murphy et al., 2022). The ABC Method also showed a high sensitivity (88.34%; 95% CI: 84.13%, 92.55%) in that stratum. However, the ABC Method's specificity was extremely low (26.99%; 95% CI: 22.58%, 31.40%); the Lasso model's specificity was also quite low, but still significantly (p < 0.05) greater than the ABC Method's (40.87%; 95% CI: 35.99%, 45.75%) (Murphy et al., 2022).

A reduced model consisting of only the Lasso model features that were significantly associated with precursor lesions (age, gender, and HP 1564) had a highly similar AUC to the full Lasso model but was only suggestive of being greater than the AUC of the ABC Method when adjusted for age and gender. Clearly, adding age and gender significantly improves both models' discrimination capability. It is possible that the *H. pylori* biomarkers that predict gastric adenocarcinoma in a specific population do not predict its precursor lesions well. However, this would be surprising given that precursors like intestinal metaplasia and dysplasia have been established as part of the common cascade towards non-cardia gastric adenocarcinoma, the predominant type of gastric cancer in this high-risk East Asian population (Correa et al., 2012).

Strengths and Limitations

A strength of this study was the detailed individual-level serum data it collected on antibody response to *H. pylori* proteins and pepsinogen levels. Furthermore, outcome categories were clinically ascertained, which should ensure a high level of validity in our classification estimates. Comparing precancerous lesions to superficial gastritis/mild CAG may have given a more valid estimate of the specificity of the test than comparing lesions to normal gastric mucosa. This is because individuals with gastritis might have a higher risk of precancerous lesions than individuals with normal stomachs (Correa et al., 2012).

The application of this Lasso-derived model was limited by the relatively small sample size of the Linqu County Study, resulting in less precise estimates. In particular, the sensitivity and specificity for the Lasso model had wide 95% confidence intervals. Additionally, the Linqu County study's design was cross-sectional, which means we cannot assess the risk of precancerous lesions in any timeframe, e.g. 10-year risk. The lack of time-to-event data meant that we were unable to fit a Cox model, which was the method used to fit the risk stratification model in the derivation data set; instead, we fit a logistic model. Nevertheless, it is promising that the Lasso model maintained a similar AUC and sensitivity to what it achieved in the derivation data set. Moreover the hazard was linear over time and, therefore, a logistic model would likely have had similar results to the Cox model (Murphy et al., 2022). Most crucially, the multiplex serology technology used to measure host response to *H. pylori* proteins in this study is not readily available in clinical laboratories in East Asia. Technology transfer from research laboratories would need to happen before the Lasso model could be implemented in clinics.

In conclusion, Linqu County is located in Shandong Province in eastern China and has an extremely high

incidence of gastric cancer (Wong et al., 2012). A non-invasive risk stratification tool could therefore be of considerable public health benefit in this region. The Lasso-generated model for gastric cancer risk stratification, consisting of gender, age, HP 0305, HP 1564, UreA, and serologically defined chronic atrophic gastritis, performed similarly in a data set of precancerous lesions to how it did in its derivation set. This suggests that it generalizes well to other populations in East Asia and to precancerous lesions which may be earlier steps on the path to carcinogenesis. The Lasso model displayed a greater specificity than the ABC Method, an existing risk stratification model currently in use in parts of East Asia. However, adding age and gender to the ABC Method substantially improved its discrimination ability and made the differences in discrimination metrics from the Lasso model non-significant. These results suggest that it is valuable to include age and gender in classification models of gastric cancer precursor lesions. Additionally, following necessary technology transfer and training, the Lasso model could be a useful addition to the landscape of general population-level screening for gastric cancer and, because these lesions are treatable, possibly improve gastric cancer interception and prevention in high-risk regions.

Data Availability Statement

Author Contribution Statement

Acknowledgements

References

Allemani C, Weir HK, Carreira H, et al (2015). Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet*, **385**, 977-1010.

Cai H, Ye F, Michel A, et al (2016). Helicobacter pylori blood biomarker for gastric cancer risk in East Asia. *Int J Epidemiol*, **45**, 774-81.

Cole P, Morrison AS (1980). Basic issues in population screening for cancer. *J Natl Cancer Inst*, **64**, 1263-72.

Colquhoun A, Arnold M, Ferlay J, et al (2015). Global patterns of cardia and non-cardia gastric cancer incidence in 2012. *Gut*, **64**, 1881-8.

Correa P, Houghton J (2007). Carcinogenesis of Helicobacter pylori. *Gastroenterology*, **133**, 659-72.

Correa P, Piazuelo MB (2012). The gastric precancerous cascade. *J Dig Dis*, **13**, 2-9.

Correa P, Piazuelo MB, Wilson KT (2010). Pathology of gastric intestinal metaplasia: clinical implications. *Am J Gastroenterol*, **105**, 493.

Dixon MF, Genta RM, Yardley JH, et al (1996). Classification and grading of gastritis: the updated Sydney system. *Am J Surg Pathol*, **20**, 1161-81.

Doorakkers E, Lagergren J, Engstrand L, et al (2016). Eradication of Helicobacter pylori and gastric cancer: a systematic review and meta-analysis of cohort studies. *J Natl Cancer Inst*, **108**.

Epplein M, Butt J, Zhang Y, et al (2018). Validation of a Blood Biomarker for Identification of Individuals at High Risk for Gastric CancerValidation of a Gastric Cancer Risk Biomarker. *Cancer Epidemiol Biomarkers Prev*, **27**, 1472-79.

Ferlay J, Soerjomataram I, Dikshit R, et al (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*, **136**, E359-E86.

Friedman J, Hastie T, Tibshirani R (2010). Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*, **33**, 1.

Li B, Taylor PR, Li JY, et al (1993). Linxian nutrition intervention trials design, methods, participant characteristics, and compliance. *Ann Epidemiol*, **3**, 577-85.

Liu Y, Chen PHC, Krause J, et al (2019). How to read articles that use machine learning: users' guides to the medical literature. *JAMA*, **322**, 1806-16.

Michel A, Waterboer T, Kist M, et al (2009). Helicobacter pylori multiplex serology. *Helicobacter*, **14**, 525-35.

Murphy JD, Olshan AF, Lin F-C, et al (2022). A Predictive Model of Noncardia Gastric Adenocarcinoma Risk Using Antibody Response to Helicobacter pylori Proteins and Pepsinogen. *Cancer Epidemiol Biomarkers Prev*, **31**, 811-20.

Robin X, Turck N, Hainard A, et al (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform*, **12**, 1-8.

RStudio T (2020). RStudio: integrated development for R. Rstudio Team, PBC, Boston, MA URL http://www. rstudio. com.

Rugge M, Correa P, Dixon MF, et al (2000). Gastric dysplasia: the Padova international classification. *Am J Surg Pathol*,

24, 167-76.

Sasazuki S, Inoue M, Iwasaki M, et al (2006). Effect of Helicobacter pylori infection combined with CagA and pepsinogen status on gastric cancer development among Japanese men and women: a nested case-control study. *Cancer Epidemiol Biomarkers Prev*, **15**, 1341-47.

Sugano K (2019). Effect of Helicobacter pylori eradication on the incidence of gastric cancer: a systematic review and meta-analysis. *Gastric Cancer*, **22**, 435-45.

Sung H, Ferlay J, Siegel RL, et al (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, **71**, 209-49.

Sutton P, Mitchell H (2010). Helicobacter pylori in the 21st Century (Vol. 17): CABI.

Team RC (2016). R Core Team R. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria; 2016. In: ISBN 3–900051–07–0, URL http://www. R–project. org/. Available: http://www.

Tibshirani R (1996). Regression shrinkage and selection via the lasso. *J R Stat Soc B*, **58**, 267-88.

Tibshirani R (1997). The lasso method for variable selection in the Cox model. *Stat Med*, **16**, 385-95.

Wong BC, Zhang L, Ma Jl, et al (2012). Effects of selective COX-2 inhibitor and Helicobacter pylori eradication on precancerous gastric lesions. *Gut*, **61**, 812-8.

Yamaguchi Y, Nagata Y, Hiratsuka R, et al (2016). Gastric cancer screening by combined assay for serum anti-Helicobacter pylori IgG antibody and serum pepsinogen levels-the ABC method. *Digestion*, **93**, 13-18.