# Risk Score using Demographic and Clinical Risk Factors Predicts Gastric Intestinal Metaplasia Risk in a U.S. Population

**Mimi C. Tan, MD, MPH**[1], **Quynh Ho**[2], **Theresa H. Nguyen, MD**[1,3], **Yan Liu, MS**[1,3], **Hashem B. El-Serag, MD, MPH**[1,3], **Aaron P. Thrift, PhD**[4,5]

[1]Section of Gastroenterology and Hepatology, Department of Medicine, Baylor College of Medicine, Houston, Texas, USA

[2]Texas Tech University, Lubbock, Texas, USA

[3]Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center, Houston, Texas, USA

[4]Section of Epidemiology and Population Sciences, Department of Medicine, Baylor College of Medicine, Houston, Texas, USA

[5]Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas, USA

## Abstract

**Background/Aims:** Screening for gastric intestinal metaplasia (GIM) may lead to early gastric cancer detection. We developed and validated a pre-endoscopy risk prediction model for detection of GIM based on patient-level risk factors in a U.S. population.

**Correspondence**: Mimi C. Tan, MD, MPH, Baylor College of Medicine, One Baylor Plaza, MS: BCM 285, Houston, TX 77030-3498. 713-798-0950. mc2@bcm.edu.

Authorship Statement:

  i.     Guarantor of the article: Mimi C. Tan, MD, MPH

  ii.    Specific author contributions:

         APT, HBE, MCT: analysed the data, designed the research study, wrote the paper, and contributed to the design of the study

         YL: analysed the data, wrote the paper, and contributed to the design of the study THN: collected the data, wrote the paper

         QH: wrote the paper

  iii.   All authors have approved the final version of the manuscript.

  iv.    All work was independent of the funding source.

**Ethics approval**: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the Institutional Review Boards for Baylor College of Medicine and the MEDVAMC.

**Consent to participate:** Informed consent was obtained from all individual participants included in the study.

**Conflicts of interest:** The authors report no competing interests for this publication.

**Methods:** We used data from 423 GIM cases and 1796 controls from a cross-sectional study among primary care and endoscopy clinic patients at the Houston VA. We developed the model using backwards stepwise regression and assessed discrimination using area under the receiver operating characteristic (AUROC). The model was internally validated using cross-validation and bootstrapping. The final expanded model was compared to a model including *H. pylori* infection alone and a baseline model including remaining terms without *H. pylori*.

**Results:** Male sex, older age, non-White race/ethnicity, smoking status, and *H. pylori* were associated with GIM risk. The expanded model including these terms had AUROC 0.73 (95%CI 0.71–0.76) for predicting GIM and AUROC 0.82 (95%CI 0.79–0.86) for extensive GIM. This model discriminated better than a model including only *H. pylori* (AUROC 0.66; 95%CI 0.63–0.68) and the baseline model (AUROC 0.67; 95%CI 0.64–0.70). The expanded model performed similarly among primary care (AUROC 0.75) and endoscopy (AUROC 0.73) patients. The expanded model showed sufficient internal validity (cross-validation AUROC 0.72) with little evidence of over-fitting.

**Conclusions:** We develop and validated a non-invasive risk model for GIM detection in a U.S. population that included terms for sex, age, race/ethnicity, smoking status, and *H. pylori* infection. Validated risk models would identify individuals with GIM who should be referred for endoscopic screening.

## Introduction

Non-cardia gastric adenocarcinoma (hereafter referred to as gastric cancer) is the third leading cause of cancer-related death worldwide[1, 2]. Rates of gastric cancer are increasing among adults <50 years old in the United States (U.S.), especially among Hispanics[3, 4]. Gastric intestinal metaplasia (GIM), a precursor lesion of gastric cancer[5], may be an ideal target for early identification in a gastric cancer screening program. However, because GIM is an asymptomatic condition, those at highest risk for GIM from the general population need to be identified and referred for upper endoscopy with gastric biopsies[6]. Currently there is not a screening strategy to identify these high-risk patients with GIM in the U.S.

Countries with the highest incidence rates of gastric cancer (i.e., Korea and Japan) have a universal gastric cancer screening strategy using only age cut-off of 40 years[7]. However, screening all adults in the U.S. aged 40 years is not cost effective given the overall lower incidence of gastric cancer[8]. Meanwhile, several populations in the U.S. (i.e., Hispanics, first generation Asians) have gastric cancer incidence rates similar to those in Asia (age-adjusted incidence 17.1–29.3 per 100,000)[9, 10]. A screening strategy based on risk stratification is needed to identify those individuals in the population at highest risk to be referred for endoscopic screening for GIM and early cancer.

Several demographic and clinical risk factors have been associated with GIM among U.S. populations. *Helicobacter pylori* is the main risk factor for GIM development as

GIM develops in the setting of inflammation and atrophy of gastric mucosa associated with *H. pylori* infection[6, 11]. Other possible risk factors among U.S. populations include male gender[12, 13], non-White race/ethnicity[12, 14, 15], older age[14, 16], and smoking[17, 18]. A predictive model that combines demographic and clinical risk factors may identify individuals in the general population at high risk of GIM. The aim of this study was to develop and internally validate a comprehensive risk prediction model for GIM based on well-described demographic, clinical and lifestyle risk factors in the U.S. population.

## Methods

We developed a risk prediction model for GIM using data from a cross-sectional study of U.S. patients at the Michael E. DeBakey VA Medical Center (MEDVAMC) in Houston, Texas. Eligible patients were identified from primary care and endoscopy clinics and invited to participate in the research study. Those who provided written informed consent answered questionnaires and underwent an esophagogastroduodenoscopy (EGD) with gastric biopsies as part of the research study.

We have previously described the study population and procedures[12, 19]. Briefly, patients were recruited from two sources: 1) consecutive patients from 1 of 7 primary care clinics who were eligible for average risk colon cancer screening colonoscopy and gave consent to additionally undergo EGD, and 2) consecutive patients previously scheduled for an EGD due to gastrointestinal symptoms and gave consent for additional gastric biopsies as part of the research study. We chose to randomly sample from these 2 groups as they are representative of the entire VA population comprising asymptomatic individuals undergoing screening (i.e., primary care population) and symptomatic individuals undergoing EGD due to gastrointestinal symptoms (i.e., endoscopy clinic population). The two populations therefore represent the source population at MEDVAMC for GIM. We recruited patients age 40–80 years (50–80 years in the primary care group due to colon cancer screening commencing at age 50) and excluded patients with: 1) previous gastroesophageal surgery; 2) previous cancer; 3) use of anticoagulants; 4) platelet counts <70,000, ascites, or gastroesophageal varices; or 5) history of major stroke or mental condition inhibiting interview ability. In the primary care group, 43% of invited eligible patients completed the study (i.e., completed study questionnaire and study EGD). In the endoscopy clinic group, 70% of those who were invited to participate completed the study. The study was approved by the Institutional Review Boards for Baylor College of Medicine and the MEDVAMC.

All patients completed questionnaires that ascertained age, sex, race/ethnicity, history of alcohol consumption and smoking usage, medical history, gastroesophageal reflux disease (GERD) symptoms, and use of proton pump inhibitors (PPI), histamine-2 receptor antagonists (H2RA), aspirin, and nonsteroidal anti-inflammatory drugs (NSAIDs). Each participant's anthropometric measurements were recorded (i.e., weight, height, and waist and hip circumference). Body mass index (BMI) and waist-to-hip ratio (WHR) were calculated. WHR was categorized as either high ( 0.9 for men, 0.85 for women) or low.

As part of the study, all patients underwent EGD with gastric mapping biopsies consisting of 2 biopsies from each location (i.e., antrum [both greater and lesser curvature], corpus [proximal greater curvature, proximal lesser curvature, with optional additional biopsies at distal greater curvature and distal lesser curvature], incisura, and cardia) according to the adoption of the Sydney System[20]. Biopsy specimens were embedded in paraffin, oriented on edge, sectioned in 5-μ sections, and stained with hematoxylin and eosin, alcian blue at pH 2.5; and in case of negative staining for *H. pylori*, a modified silver stain; and alcian blue–periodic acid Schiff stain. The presence and severity of histopathological findings including GIM and *H. pylori* was independently determined by two blinded gastrointestinal pathologists. Disagreements in pathology reads were determined by a third pathologist.

We defined GIM cases as those with GIM on at least 1 non-cardia gastric biopsy. Controls were participants without GIM on all non-cardia gastric biopsies. Extensive GIM was defined as presence of GIM in both antrum and corpus biopsies. Patients were considered to have *H. pylori* infection if *H. pylori* organisms were found on histopathology of 1 gastric biopsy site or isolated on gastric tissue culture. To process cultures for *H. pylori*, frozen tissue specimens were thawed, homogenized, and inoculated onto Brain Heart Infusion (nutrient rich agar ideal for culturing fastidious microorganisms) and *H. pylori* Special Peptone Agar plates with 7% horse blood. The plates were incubated at 37°C under micro-aerophilic conditions (5% O2, 10% CO2, and 85% N2) in an Anoxomat jar for up to two weeks. Positive growth was transferred to a fresh, nonselective Brain Heart Infusion blood agar plate and incubated for 48–72 hours. *H. pylori* were identified when the oxidase, catalase, and urease reactions were positive with a compatible Gram stain. To obtain a pure culture, we selected and subcultured several small round colonies from each patient's plate. Isolated strains were then stored at 80°C in cysteine storage medium containing 20% glycerol.

## Statistical Analysis

Candidate predictor variables were selected *a priori* from the literature and included: age (years), sex, race/ethnicity (non-Hispanic white, Hispanic, African American, other), BMI (<25, 25–30, 30 kg/m²), WHR (high, low), smoking status (never, former, current), alcohol status (never, former, current), presence of GERD symptoms ever, *H. pylori* infection, use of PPI/H2RA, and use of aspirin/NSAID. For comparisons between cases and controls, we used chi-square tests for categorical variables and Student t-test for continuous variables. All analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC), and a 2-tailed p-value of $< 0.05$ was considered statistically significant.

## Risk model development

We included in the multivariate model variables that were associated with GIM with $p<0.20$ in univariate analyses and performed a backwards stepwise logistic regression whereby factors losing their significance (i.e., p 0.05) in the multivariate analysis were dropped.

## Assessment of model performance

The performance of the model was assessed using tests for discrimination and calibration. We evaluated predictive discrimination using the area under the receiver operating

characteristic curve (AUROC; [also known as the *c*-statistic]) and its 95% confidence interval (95% CI). To be useful, the predicted risks must discriminate well between those with GIM (cases) and those without GIM (controls). The AUROC gives the probability that for any randomly selected pair of individuals, one case and one control, the model assigns a higher probability to the case. An AUROC of 0.5 indicates that the model has a predictive discrimination no better than chance, whereas an AUROC of 1.0 indicates a perfectly discriminating model.

AUROC was also calculated with repeated 10-fold cross-validation and bootstrapping techniques used for internal validation of the risk models. The 10-fold cross-validation splits the sample into 10 subsets randomly without replacement; one subset is used as the validation dataset while the remaining 9 are used as the training dataset in the logistic regression model to calculate the predicted probability for each validation observation. This is repeated 9 more times then the average AUROC is calculated using the 10 runs. The bootstrap method first generates 100 bootstrap samples with replacement from the original sample. On each bootstrap sample, the performance was measured for the bootstrap sample and the original sample, and the average difference between the two performances forms an estimate of the optimism. The overall AUROC is the average of the AUROC for all iterations.

The predictive ability of our final model (referred to as expanded model) for detecting GIM was compared with: 1) a model including *H. pylori* infection alone and 2) a baseline model that included the demographic and clinical risk factors without *H. pylori* infection. We additionally examined the model performance stratified by the recruitment source to determine robustness of prediction within an asymptomatic primary care population and a symptomatic cohort previously scheduled for endoscopy. We lastly determined predictive ability of all 3 models for only extensive GIM, defined as GIM in both antrum and corpus biopsies, which is associated with higher cancer risk than overall GIM. The second measure calculated was calibration, which compares the predicted probabilities with the observed risk. We evaluated calibration using the Hosmer-Lemeshow goodness-of-fit statistic[21] where a high p-value (>0.05) indicates that the model is well calibrated.

We additionally calculated the Youden J statistic (index) which selects the optimal predicted probability cut-off of the model based on the maximum vertical distance between the ROC curve and diagonal line. The goal is to maximize the difference between true positive and false positive. We reported the corresponding sensitivity and specificity associated with the optimal cut-off calculated by the Youden index and for varying probability thresholds for the expanded model.

## Results

Table 1 shows select characteristics of the 423 cases with GIM and 1796 controls without GIM with a calculated prevalence of 19.1% overall (21.3% among primary care patients; 18.3% among endoscopy patients). Patients with GIM were on average older (62.1 vs. 59.9; p<0.001) and more likely to be male (97.2% vs. 90.8%; p<0.001) and of non-White race/ethnicity (58.6% vs 39.0%; p<0.001). Overall, 219 cases (51.8%) and 394 controls (21.9%)

had *H. pylori* infection; of 1407 patients who additionally had *H. pylori* culture performed, 25.2% had *H. pylori* infection. Omeprazole was the most commonly used PPI in 41.6% of 176 cases and 43.8% of 787 controls. Among omeprazole users, 95.7% were daily users (94.3% of cases, 96.1% of controls).

We examined three models: 1) a model that included only *H. pylori*, 2) baseline model with sex, age, race/ethnicity, and smoking status, and 3) an expanded model that combined *H. pylori* and the baseline model. In the expanded model, the following risk factors were significantly associated with risk of GIM: male sex (OR 2.64; 95% CI 1.38–5.03), age (ref <60 years; 60–69.9 years: OR 1.69; 95% CI 1.29–2.21;   70 years: OR 2.44; 95% CI 1.66–3.59), non-White race/ethnicity (ref White; African-American: OR 1.86; 95% CI 1.43–2.43; Hispanic: OR 2.32; 95% CI 1.61–3.34;), smoking status (ref never smoker; current smoker: OR 2.06; 95% CI 1.48–2.86; former smoker: OR 1.37; 95% CI 1.00–1.87;), and *H. pylori* infection (OR 3.57; 95% CI 2.79–4.55) (Table 2).

### Risk model performance

The expanded model provided highest discriminatory ability among the 3 models tested. The AUROC for the *H. pylori* only model was 0.66 (95% CI 0.63–0.68) and for the baseline model was 0.67 (95% CI 0.64–0.70). When *H. pylori* was added to the baseline model to form the expanded model, the AUROC increased to 0.73 (95% CI 0.71–0.76) (Figure 1; Table 3). The expanded model showed good internal validity and we found little evidence of over-fitting (10-fold cross validation AUROC 0.72; Bootstrapping corrected AUROC 0.72; compared with AUROC 0.73 in the study dataset). Finally, the expanded model was well-calibrated according to the Hosmer-Lemeshow test (p=0.29).

The expanded model performed equally well among the 1655 endoscopy clinic patients (AUROC 0.73; 95% CI 0.70–0.77) and 564 primary care clinic patients (AUROC 0.75; 95% CI 0.70–0.80) (Supplementary Figure 1). Furthermore, among 108 patients with extensive GIM compared to 1796 controls, the expanded model had higher discrimination ability to predict extensive GIM (AUROC 0.82; 95% CI 0.79–0.86) than the *H. pylori* only model (AUROC 0.78; 95 CI 0.64–0.73) or the baseline model (AUROC 0.78; 95% CI 0.74–0.82) (Figure 2).

For the expanded model, the Youden's J index showed the optimal cut-off to have sensitivity 72.7% and specificity 65.8%. To assess the potential effects of using the expanded model to guide referral for endoscopic screening for GIM, we calculated the proportion of patients that would be referred for endoscopy at different probability thresholds for GIM and the associated sensitivities and specificities (Table 4). The first row gives the scenario of referring every patient with GIM for endoscopy and therefore identifying all patients who have GIM (i.e., sensitivity is 100%). If patients are referred for endoscopy only if their predicted probability of GIM is, for example, 30% or more, the proportion of patients referred for endoscopy will be reduced to 21.7%. At that threshold however, about 55.9% of GIM cases will not be referred for endoscopy (sensitivity, 44.1%). As the threshold increases, the number of referrals is reduced, however, the number of patients with GIM who will not be referred for endoscopy increases.

## Discussion

Using data from a large, cross-sectional study among U.S. veterans, we developed and internally validated a comprehensive risk prediction model for GIM. The final "expanded model" included terms for sex, age, race/ethnicity, smoking status, and *H. pylori* infection and therefore can be used pre-endoscopy using commonly and easily obtainable data. Internal validation approaches showed that this expanded model was well-calibrated and performed well in discriminating between patients with GIM and controls without GIM (AUROC=0.73) and even better at discriminating patients with extensive GIM (AUROC=0.82). The expanded model performed better, in terms of model discrimination, than did the *H. pylori* only model (AUROC=0.66) and the baseline model (terms for sex, age, race/ethnicity, and smoking status; AUROC=0.67).

Applying this model for making a decision has the potential to improve screening and detection of GIM. Based on the Youden index, the optimal cut-off of the expanded model had a sensitivity 72.7% and specificity 65.8%. At this cut-off, one would detect 73% of GIM cases, while missing 27%; for a population with 20% prevalence of GIM, for every 35 true positive EGDs, there would be 65 unnecessary EGDs. In general, determining an acceptable threshold involves a trade-off between sensitivity and specificity. In screening for a lethal cancer for example, high sensitivity is desirable, whereas for diseases with lower severity, a lower sensitivity can be tolerated. For our GIM model, if the model is primarily used as a first-step in GIM screening of a high-risk population, we could sacrifice specificity for a higher sensitivity (92.3%; specificity 14.2%). This would ensure that almost all patients with GIM would be captured and referred for confirmatory endoscopy although the yield of finding GIM on endoscopy would be low. Given the prevalence of GIM in our study cohort (19%), with sensitivity 92.3%, specificity 14.2% of the risk model, for every 18 true positives, we would perform unnecessary endoscopy in 69 (false positive) and miss 1 GIM case (false negative). While these estimates may sound discouraging, they compare well with other screening tests for example (breast cancer screening mammography sensitivity 97%, specificity 64.5%[22]; lung cancer screening chest computed tomography [CT] sensitivity 94%, specificity 73%[23]).

Rather than specifying screening among persons with one or multiple unweighted risk factors, we propose screening according to individual probability for having GIM according to the output of our expanded model that weighs the presence/absence of multiple factors based on their magnitude of association with GIM risk. The optimum threshold will likely be determined by acceptable sensitivity and specificity estimates and may vary depending on population or clinical setting. This study is a first step towards the goal of developing a clinically useable tool for predicting the presence of GIM using measurable, pre-endoscopy risk factors. Such risk prediction tools have been developed to estimate liver cancer risk[24].One crucial determining factor of eventual acceptability and applicability of any screening strategy is cost effectiveness. These probabilities provide essential information for cost effectiveness analyses that will further inform decision making. Improvements in the predictive ability of this model are required and could be achieved in two broad approaches. One approach is adding or optimizing the measurement of the pre-endoscopy variables. The other approach is to combine with predictive models used to screen other upper GI

conditions (e.g., Barrett's esophagus) and arrive at one predictive model used in screening for all upper GI cancers.

PPI use was not a significant predictor of GIM in our study. While we did not have granular data on PPI dosage and duration, we found that most PPI users were daily users and there was not a significant difference in PPI use between GIM cases and controls. PPIs have been associated in previous studies with an increased risk of GIM[25] and possibly gastric cancer[26] in *H. pylori*-positive patients, especially with longer duration of use. Our study found traditional risk factors (i.e., *H. pylori*, non-white race/ethnicity) to be predictive of GIM rather than PPI use, which has a less definitive association with GIM.

To our knowledge, ours is the first study using demographic and clinical risk factors to develop a GIM risk model to direct endoscopic screening for GIM in the U.S. Previous studies have used screening biomarkers (e.g., pepsinogens) for detection of atrophic gastritis and GIM in high risk countries. However, pepsinogen tests have a wide range in reported sensitivity (25%–91%)[27] and specificity (39%–100%)[27], depending on cut-off value used and population prevalence of disease, but is generally thought to be between 70%–80%[27–29]. Apart from additional cost and spotty availability, pepsinogen testing is also susceptible to reduced sensitivity in the common setting of PPI use due to reduced intragastric acidity resulting in increased pepsinogens[30]. Moreover, *H. pylori* infection raises pepsinogen levels related to cytokines and inflammation[31], which may result in a falsely elevated pepsinogen 1. Therefore, the usefulness of pepsinogen testing is unclear[32].

Strengths of this study include the prospective recruitment of patients who underwent systematic gastric mapping biopsies without regard to gastrointestinal symptoms or endoscopic findings, thereby minimizing selection bias for case and control determination. Additionally, our study cohort included a mix of asymptomatic patients recruited for this study and patients undergoing endoscopy due to gastrointestinal symptoms, thus representing the total population of interest. Any risk prediction model to be used in population screening must be developed in a sample population representative of the overall population. Not only was our risk prediction model developed in a representative population, it was furthermore found to perform equally well among asymptomatic primary care clinic patients and symptomatic endoscopy clinic patients. We additionally employed multiple methods of internal validation, including cross validation and bootstrapping validation.

There are few specific areas for further improvements in the predictive ability of our model. *H. pylori* was tested in our model using biopsy-based tests (i.e., histopathology, culture), and while we believe that there will be high concordance with results of non-invasive tests (e.g., urea breath test, serology), this needs to be examined and may improve the sensitivity of the model since serology would account for both past and present infections. Only 8% of our cohort were women, and therefore, the parameter estimate for sex may not be stable and may require further optimization in future studies with a higher proportion of women. There were also very few patients born or raised in southeast Asia where GIM and gastric cancer are very common, so this factor could not be examined. Other information that may prove important include findings of previous endoscopy (e.g., absence of GIM is likely to be a strong negative predictive value), family history of gastric cancer, or blood based

biomarkers. This model requires externally validation and likely further optimization among different cohorts (geographic validation) that include more women, younger, and immigrant populations.

In summary, we developed and internally validated a risk prediction model which estimates the likelihood of having GIM in both primary care patients and patients being referred for upper endoscopy. The final model, which included terms for sex, age, race/ethnicity, smoking status, and *H. pylori* status, outperformed a model including only a term for *H. pylori* infection. Such a risk stratification model could be used as a screening tool to identify patients at high risk of GIM in the U.S. to refer for endoscopic screening. This model, however, could be considered as a starting point for further development and validation, as a number of risk factors were not included and genetic information may also be important in predicting risk of GIM. The inclusion of other environmental risk factors and the extension of the model to include biomarkers may go further to improving performance.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Grant support:

## References

1. Bray F, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424. [PubMed: 30207593]

2. Thrift AP, et al. Burden of Gastric Cancer. Clin Gastroenterol Hepatol 2020;18:534–542. [PubMed: 31362118]

3. Wang Z, et al. Incidence of gastric cancer in the USA during 1999 to 2013: a 50-state analysis. Int J Epidemiol 2018.

4. Wang Z, et al. Increasing Incidence of Advanced Non-cardia Gastric Cancers Among Younger Hispanics in the USA. Dig Dis Sci 2020.

5. Correa P, et al. A model for gastric cancer epidemiology. Lancet 1975;2:58–60. [PubMed: 49653]

6. Asaka M, et al. Atrophic changes of gastric mucosa are caused by Helicobacter pylori infection rather than aging: studies in asymptomatic Japanese adults. Helicobacter 1996;1:52–6. [PubMed: 9398913]

7. Kim GH, et al. Screening and surveillance for gastric cancer in the United States: Is it needed? Gastrointest Endosc 2016;84:18–28. [PubMed: 26940296]

8. Saumoy M, et al. Cost Effectiveness of Gastric Cancer Screening According to Race and Ethnicity. Gastroenterology 2018;155:648–660. [PubMed: 29778607]

9. Lui FH, et al. Ethnic disparities in gastric cancer incidence and survival in the USA: an updated analysis of 1992–2009 SEER data. Dig Dis Sci 2014;59:3027–34. [PubMed: 25030941]

10. Miller BA, et al. Cancer incidence and mortality patterns among specific Asian and Pacific Islander populations in the U.S. Cancer Causes Control 2008;19:227–56. [PubMed: 18066673]

11. IARC Helicobacter pylori Working Group. Helicobacter pylori eradication as a strategy for preventing gastric cancer (IARC Working Group Reports, No. 8). Lyon, 2014.

12. Tan MC, et al. Demographic and Lifestyle Risk Factors for Gastric Intestinal Metaplasia Among US Veterans. Am J Gastroenterol 2020;115:381–387. [PubMed: 31899705]

13. Woo Y, et al. Screening endoscopy finds high prevalence of Helicobacter pylori and intestinal metaplasia in Korean American with limited access to health care. J Surg Oncol 2017;116:172–176. [PubMed: 28445591]

14. Da B, et al. High-risk symptoms do not predict gastric cancer precursors. Helicobacter 2019;24:e12548. [PubMed: 30412322]

15. Nguyen TH, et al. Prevalence of Gastric Intestinal Metaplasia in a Multiethnic US Veterans Population. Clin Gastroenterol Hepatol 2020;14:S1542–3565(20)30325–6. Online ahead of print.

16. Oh S, et al. Risk factors of atrophic gastritis and intestinal metaplasia in first-degree relatives of gastric cancer patients compared with age-sex matched controls. J Cancer Prev 2013;18:149–60. [PubMed: 25337541]

17. Stemmermann GN, et al. Impact of diet and smoking on risk of developing intestinal metaplasia of the stomach. Dig Dis Sci 1990;35:433–8. [PubMed: 2318088]

18. Fontham ET, et al. Determinants of Helicobacter pylori infection and chronic gastritis. Am J Gastroenterol 1995;90:1094–101. [PubMed: 7611204]

19. Fischbach LA, et al. Association between Helicobacter pylori and Barrett's esophagus: a case-control study. Am J Gastroenterol 2014;109:357–68. [PubMed: 24419485]

20. Dixon MF, et al. Classification and grading of gastritis. The updated Sydney System. International Workshop on the Histopathology of Gastritis, Houston 1994. Am J Surg Pathol 1996;20:1161–81. [PubMed: 8827022]

21. Steyerberg EW, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology (Cambridge, Mass.) 2010;21:128–138.

22. Zeeshan M, et al. Diagnostic Accuracy of Digital Mammography in the Detection of Breast Cancer. Cureus 2018;10:e2448. [PubMed: 29888152]

23. Church TR, et al. Results of initial low-dose computed tomographic screening for lung cancer. N Engl J Med 2013;368:1980–91. [PubMed: 23697514]

24. Tayob N, et al. Validation of the Hepatocellular Carcinoma Early Detection Screening (HES) Algorithm in a Cohort of Veterans With Cirrhosis. Clin Gastroenterol Hepatol 2019;17:1886–1893.e5. [PubMed: 30557738]

25. Snir Y, et al. Dose-dependent association of proton pump inhibitors use with gastric intestinal metaplasia among Helicobacter pylori-positive patients. United European Gastroenterol J 2021;9:343–353.

26. Cheung KS, et al. Long-term proton pump inhibitors and risk of gastric cancer development after treatment for Helicobacter pylori: a population-based study. Gut 2018;67:28–35. [PubMed: 29089382]

27. Huang YK, et al. Significance of Serum Pepsinogens as a Biomarker for Gastric Cancer and Atrophic Gastritis Screening: A Systematic Review and Meta-Analysis. PLoS One 2015;10:e0142080. [PubMed: 26556485]

28. Dinis-Ribeiro M, et al. Meta-analysis on the validity of pepsinogen test for gastric carcinoma, dysplasia or chronic atrophic gastritis screening. J Med Screen 2004;11:141–7. [PubMed: 15333273]

29. Dinis-Ribeiro M, et al. Validity of serum pepsinogen I/II ratio for the diagnosis of gastric epithelial dysplasia and intestinal metaplasia during the follow-up of patients at risk for intestinal-type gastric adenocarcinoma. Neoplasia 2004;6:449–56. [PubMed: 15548353]

30. Agréus L, et al. Clinical use of proton-pump inhibitors but not H2-blockers or antacid/alginates raises the serum levels of amidated gastrin-17, pepsinogen I and pepsinogen II in a random adult population. Scand J Gastroenterol 2009;44:564–70. [PubMed: 19263272]

31. Lorente S, et al. Helicobacter pylori stimulates pepsinogen secretion from isolated human peptic cells. Gut 2002;50:13–8. [PubMed: 11772960]

32. Tan MC, et al. Gastric cancer risk stratification and surveillance after Helicobacter pylori eradication: 2020. Gastrointest Endosc 2019;90:457–460. [PubMed: 31439130]
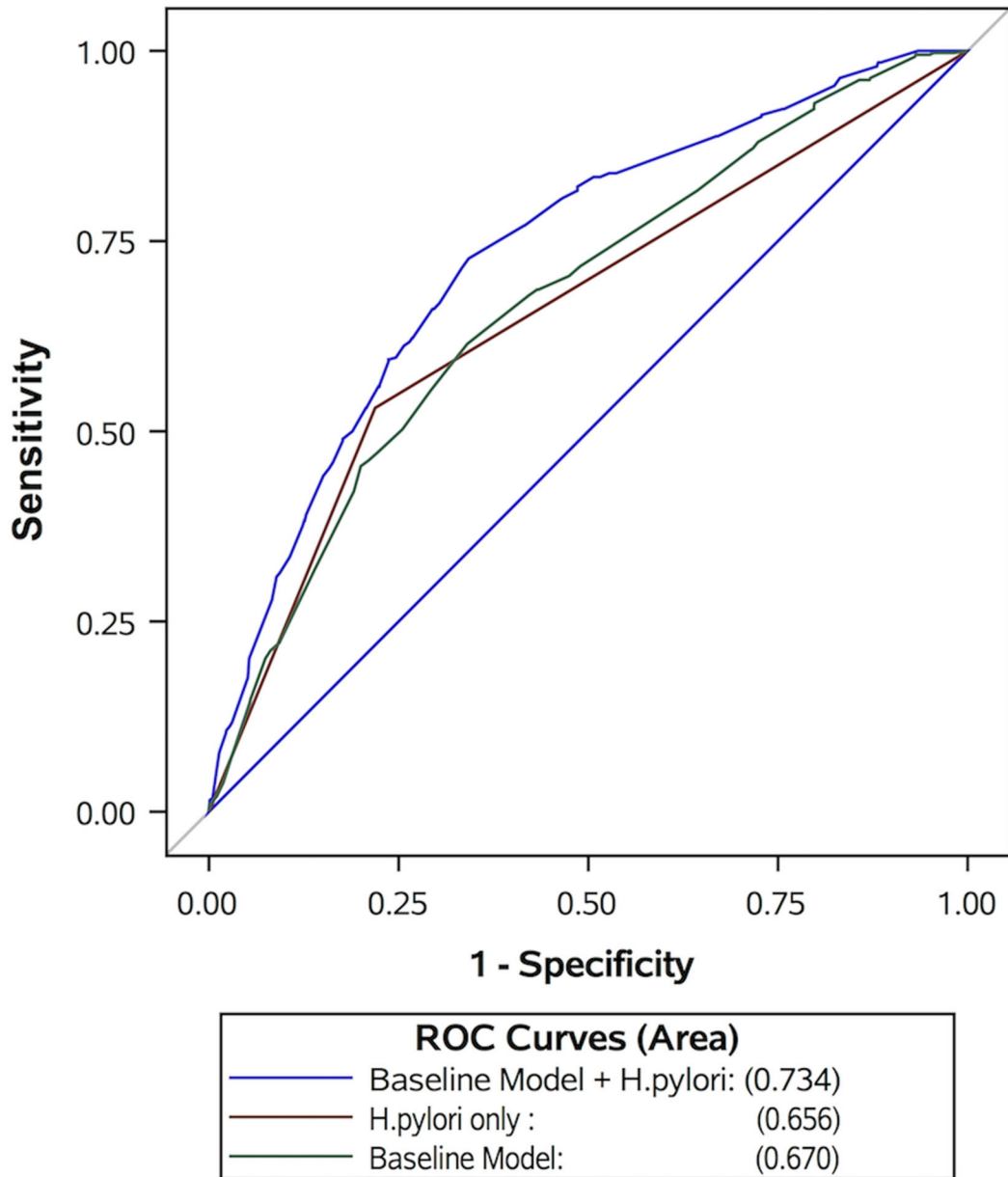
**Figure 1.**
The receiver operator characteristic (ROC) curve of 3 models in predicting gastric intestinal metaplasia: 1) *H. pylori* alone (red), 2) baseline model without *H. pylori* (green), and 3) expanded model that includes *H. pylori* and baseline model (blue).
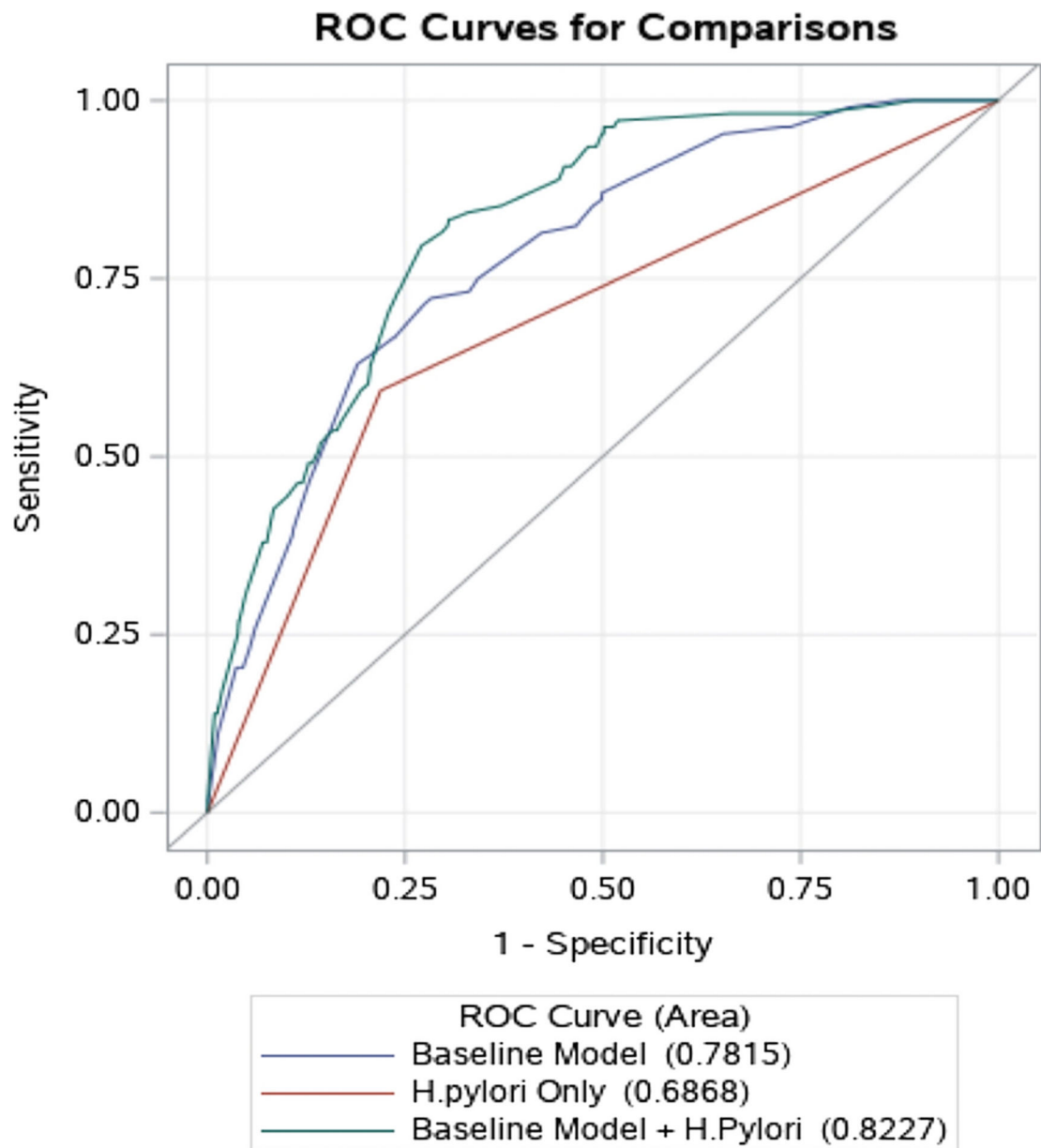
**Figure 2.**
The receiver operator characteristic (ROC) curve of 3 models in predicting extensive gastric intestinal metaplasia: 1) *H. pylori* alone (red), 2) baseline model without *H. pylori* (blue), and 3) expanded model that includes *H. pylori* and baseline model (green).

**Table 1.**

Patient and clinical characteristics of 423 cases with gastric intestinal metaplasia compared to 1796 controls from the Michael E. DeBakey VA Medical Center

|  | Cases (n=423) | Controls (n=1796) | P-value |
|---|---|---|---|
| **Recruitment source** |  |  | 0.121 |
| Endoscopy | 303 (71.63) | 1352 (75.28) |  |
| Primary Care | 120 (28.37) | 444 (24.72) |  |
| **Age** |  |  |  |
| <60 | 138 (32.62) | 761 (42.37) | <0.001 |
| 60–69 | 218 (51.54) | 842 (46.88) |  |
| 70 | 67 (15.84) | 193 (10.75) |  |
| **Sex** |  |  | <0.001 |
| Male | 411 (97.16) | 1630 (90.76) |  |
| Female | 12 (2.84) | 166 (9.24) |  |
| **Race/Ethnicity** |  |  | <.0001 |
| White | 175 (41.37) | 1095 (60.97) |  |
| Hispanic | 62 (14.66) | 148 (8.24) |  |
| Black | 178 (42.08) | 521 (29.01) |  |
| Other/Unknown | 8 (1.89) | 32 (1.78) |  |
| **BMI (kg/m2)** |  |  | 0.080 |
| <25 | 89 (21.04) | 318 (17.71) |  |
| 25–29 | 163 (38.53) | 634 (35.03) |  |
| 30 | 171 (40.43) | 840 (46.77) |  |
| Unknown/missing | 0 (0.00) | 4 (0.22) |  |
| **Waist-to-hip ratio** |  |  |  |
| Low | 60 (14.18) | 248 (13.81) | 0.712 |
| High | 349 (82.51) | 1501 (83.57) |  |
| Unknown/Missing | 14 (3.31) | 47 (2.62) |  |
| **Smoking status** |  |  | <0.001 |
| Never smoked | 82 (19.39) | 501 (27.90) |  |
| Current smoker | 138 (32.62) | 473 (26.34) |  |
| Former smoker | 185 (43.74) | 718 (39.98) |  |
| Unknown/missing | 18 (4.26) | 104 (5.79) |  |
| **Alcohol status** |  |  | 0.169 |
| Never drinker | 26 (6.15) | 151 (8.41) |  |
| Current drinker | 211 (49.88) | 891 (49.61) |  |
| Former drinker | 166 (39.24) | 639 (35.58) |  |
| Unknown/missing | 20 (4.73) | 115 (6.40) |  |
| **GERD symptoms** |  |  | 0.047 |
| No | 224 (52.96) | 832 (46.33) |  |

| | Cases (n=423) | Controls (n=1796) | P-value |
|---|---|---|---|
| Yes | 180 (42.55) | 865 (48.16) | |
| Unknown/missing | 19 (4.49) | 99 (5.51) | |
| *Helicobacter pylori* | | | <0.001 |
| No | 199 (47.04) | 1377 (76.67) | |
| Yes | 219 (51.77) | 394 (21.94) | |
| Unknown/missing | 5 (1.18) | 25 (1.39) | |
| **PPI/H2RA use** | | | 0.076 |
| No | 180 (42.55) | 659 (36.69) | |
| Yes | 218 (51.54) | 1010 (56.24) | |
| Unknown/missing | 25 (5.91) | 127 (7.07) | |
| **NSAID use** | | | 0.267 |
| No | 153 (36.17) | 717 (39.92) | |
| Less than daily | 18 (4.26) | 83 (4.62) | |
| At least daily | 188 (44.44) | 704 (39.20) | |
| Unknown/missing | 64 (15.13) | 292 (16.26) | |

BMI: body mass index; GERD: gastroesophageal reflux disease; PPI: proton pump inhibitor; H2RA: histamine-2 receptor antagonist; NSAID: non-steroidal anti-inflammatory drug

**Table 2.**

Associations of demographic and clinical risk factors with the presence of gastric intestinal metaplasia among 423 cases with and 1796 controls without gastric intestinal metaplasia represented in 3 models: 1) *H. pylori* alone, 2) baseline model without *H. pylori*, and 3) expanded model that includes *H. pylori* and baseline model.

| | *H. pylori* alone | | | Baseline Model | | | Expanded Model | | |
|---|---|---|---|---|---|---|---|---|---|
| **Effect** | **OR** | **95% (CI)** | | **OR** | **95% (CI)** | | **OR** | **95% (CI)** | |
| **Sex**: Male (Ref Female) | | | | 1.56 | 1.20 | 2.02 | 2.64 | 1.38 | 5.03 |
| **Age**: 60–69.9 (Ref <60) | | | | 2.30 | 1.58 | 3.34 | 1.69 | 1.29 | 2.21 |
| **Age**:  70 (Ref <60) | | | | 2.98 | 1.58 | 5.62 | 2.44 | 1.66 | 3.59 |
| **Race/Ethnicity**: Black (Ref White) | | | | 2.90 | 2.04 | 4.11 | 1.86 | 1.43 | 2.43 |
| **Race/Ethnicity**: Hispanic (Ref White) | | | | 2.54 | 1.98 | 3.27 | 2.32 | 1.61 | 3.34 |
| **Smoking**: Current (Ref never) | | | | 1.43 | 1.05 | 1.93 | 2.06 | 1.48 | 2.86 |
| **Smoking**: Former (Ref never) | | | | 1.96 | 1.42 | 2.69 | 1.37 | 1.00 | 1.87 |
| *H. pylori*: Yes (Ref No) | 4.03 | 3.20 | 5.07 | | | | 3.57 | 2.79 | 4.55 |

For the expanded model, Probability

$$( \text{GIM=yes} ) = \log\left(\left[\frac{p(\text{GIM} = 1)}{1 - (p(\text{GIM} = 1))}\right]\right) = -3.880 + 0.522 * (\text{age } 60 - 69.9) + 0.892 * (\text{age } \geq 70) + 0.969 * (\text{Male}) + 0.842 * (\text{Hispanic}) + 0.623 * (\text{Black}) + 0.3138 * (\text{former smoker}) + 0.721 * (\text{current smoker}) + 1.272 * (\text{H. pylori})$$

**Table 3.**

Discrimination of the models reported using area under the receiver operating characteristics (AUROC) curve, standard error (SE), and 95% confidence interval (CI) for 3 models: 1) *H. pylori* alone, 2) baseline model without *H. pylori*, and 3) expanded model that includes *H. pylori* and baseline model.

| Models | AUROC | SE | 95% CI | |
|---|---|---|---|---|
| *H. pylori* alone | 0.656 | 0.01 | 0.629 | 0.683 |
| Baseline model | 0.670 | 0.02 | 0.641 | 0.700 |
| Expanded model (baseline model + *H. pylori)* | 0.734 | 0.01 | 0.707 | 0.761 |
| **Models by recruitment source** | | | | |
| Endoscopy clinic cohort | | | | |
| *H. pylori* alone | 0.639 | 0.02 | 0.609 | 0.672 |
| Baseline model | 0.686 | 0.02 | 0.652 | 0.720 |
| Expanded model | 0.730 | 0.02 | 0.703 | 0.766 |
| Primary care clinic cohort | | | | |
| *H. pylori* alone | 0.687 | 0.03 | 0.640 | 0.740 |
| Baseline model | 0.649 | 0.03 | 0.588 | 0.711 |
| Expanded model | 0.747 | 0.03 | 0.701 | 0.803 |

**Table 4.**

Performance of various probability thresholds for the expanded model in predicting presence of gastric intestinal metaplasia.

| Probability thresholds | Sensitivity (%) | Specificity (%) | Patients undergoing endoscopy (%) |
|:---:|:---:|:---:|:---:|
| All | 100 | 0 | 100 |
| 0.10 | 88.8 | 27.2 | 75.9 |
| 0.20 | 61.7 | 73.2 | 33.6 |
| 0.30 | 44.1 | 83.6 | 21.7 |
| 0.40 | 25.8 | 91.7 | 11.7 |
| 0.50 | 10.2 | 97.8 | 3.7 |
| 0.60 | 1.5 | 99.5 | 0.7 |
| 0.70 | 0 | 100 | 0 |
| 0.80 | 0 | 100 | 0 |
| 0.90 | 0 | 100 | 0 |
| 1.00 | 0 | 100 | 0 |