A clinical risk prediction model for Barrett's esophagus

Aaron P. Thrift[1,2], Bradley J. Kendall[2,3], Nirmala Pandeya[1], Thomas L. Vaughan[4,5], and David C. Whiteman[2] for the Study of Digestive Health

[1] *School of Population Health and* [3] *Department of Medicine, The University of Queensland, Herston Road, Herston, Queensland 4006, Australia.*

[2] *Population Health Department, Queensland Institute of Medical Research, 300 Herston Road, Herston, Queensland 4006, Australia.*

[4] *Division of Public Health Sciences, Department of Epidemiology, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.*

[5] *Department of Epidemiology, University of Washington, Seattle, Washington, USA.*

Running title: Predicting risk of Barrett's esophagus

Keywords: Barrett's esophagus; esophageal adenocarcinoma; risk; prediction; model.

Corresponding author: Mr. Aaron P. Thrift, Cancer Control Laboratory, Queensland Institute of Medical Research, Locked Bag 2000 Royal Brisbane Hospital, Queensland 4029,

Australia. Tel: +61 7 33620250. Fax: +61 7 38453502. Email: aaron.thrift@qimr.edu.au.

Disclosure of potential conflicts of interest:

The authors disclose no conflicts.

Word count: 3679 (excluding references).

Total number of figures and tables: 5.

ABSTRACT (243 words)

Barrett's esophagus (BE) is the only known precursor to esophageal adenocarcinoma. As definitive diagnosis requires costly endoscopic investigation, we sought to develop a risk prediction model to aid in deciding which patients with gastroesophageal reflux (GER) symptoms to refer for endoscopic screening for BE.

The study included data from patients with incident nondysplastic BE (n=285) and endoscopy control patients with esophageal inflammatory changes without BE ("inflammation controls", n=313). We used two phases of stepwise backwards logistic regression to identify the important predictors for BE in men and women separately: firstly including all significant covariates from univariate analyses; then fitting non-significant covariates from univariate analyses to identify those effects detectable only after adjusting for other factors. The final model pooled these predictors and was externally validated for discrimination and calibration using data from a BE study conducted in western Washington State, USA.

The final risk model included terms for age, sex, smoking status, body mass index, highest level of education, and frequency of use of acid suppressant medications (area under the ROC curve, 0.70, 95%CI 0.66-0.74). The model had moderate discrimination in the external dataset (area under the ROC curve, 0.61, 95%CI 0.56-0.66). The model was well calibrated (Hosmer-Lemeshow test, p=0.75), with predicted probability and observed risk highly correlated.

The prediction model performed reasonably well and has the potential to be an effective and useful clinical tool in selecting patients with GER symptoms to refer for endoscopic screening for BE.

INTRODUCTION

Esophageal cancer is the sixth leading cause of cancer-related mortality worldwide; in 2008, an estimated 482,000 new cases and 407,000 cancer-associated deaths were predicted (1). In the UK, the US and most other Western countries, esophageal adenocarcinoma (EAC) is the predominant histological subtype of esophageal cancer (2). From 1975 to 2001, the rate of increase in incidence of EAC was among the highest for any cancer in Western countries, with incidence increasing by a factor of 6 in the US (3). In 2007, the UK reported the world's highest rate of EAC incidence (approximately 5 per 100,000) (4). Survival for EAC is poor as most EAC patients present with metastatic disease and less than 20% survive for 5 years (5).

Almost all cases of EAC are thought to arise from underlying Barrett's esophagus (BE), progressing through low to high grade dysplasia (6); however the risk of progressing from BE to EAC is unclear. Initial studies suggested that the absolute rate of progression was between 5 to 10 per 1000 person-years (7), however a recent study in Denmark reported a much lower estimate of 1.2 per 1000 person-years (8). It is widely accepted that gastroesophageal reflux (GER) is the principal underlying cause of BE (9), however a number of large population-based studies have identified other factors associated with BE, including central obesity (10, 11), and smoking (12).

The identification of patients with BE holds the greatest potential, at least in theory, for reducing mortality from EAC. The challenge remains however, to determine which patients presenting with GER symptoms are more likely to have underlying, undiagnosed BE, based on their history and symptoms. Current guidelines from both the British Society of Gastroenterology (13) and the American Gastroenterological Association (14) do not

recommend endoscopic BE screening of the general GER population. The American
Gastroenterological Association do recommend endoscopic BE screening for those with
multiple risk factors for EAC (including age 50 years or older, male sex, white race, chronic
GER, hiatal hernia, and elevated body mass index). However, these guidelines have been
developed based on expert opinion and weak evidence, and have not been validated (14, 15).

There is growing interest in developing risk prediction models to assist practitioners in
identifying patients who may benefit most from investigations and interventions. Models
have been developed and used for a variety of conditions ranging from cardiovascular
diseases (16), to cancers of the breast (17-20), lung (21-23), prostate (24), and bladder (25),
and to melanoma (26). These models have been shown to be clinically useful and reliable at
the population level, providing a cost-effective approach to disease prevention and treatment.
While models have been developed to predict BE in symptomatic patients, they have focused
on only a restricted list of factors (e.g., age, sex and upper gastrointestinal symptoms) and did
not consider phenotypic and environmental factors which may explain a high proportion of
BE cases not attributable to GER (27, 28). If a comprehensive risk model could be
developed, it could have the potential to influence clinical decision-making in the care of
patients with GER.

We aimed to develop and validate a risk prediction model for quantifying the probability of
BE in patients with frequent GER symptoms and to demonstrate the potential utility of this
model in the clinical setting for selection of patients for endoscopic screening for the
presence of BE.

MATERIAL AND METHODS

Study population

To derive the prediction model, we used data from participants in the Study of Digestive Health, a population-based study of BE and reflux-related conditions conducted in Brisbane, Australia. The study population has been described in full previously (12). We then validated the model in a separate case-control study of BE conducted in western Washington State, USA (29).

In brief, eligible cases for the Study of Digestive Health were residents of metropolitan Brisbane aged 18-79 years with a new (incident) histologically confirmed diagnosis of BE (for nondysplastic BE cases) or dysplasia (for dysplastic BE cases) between 1 February 2003 and 30 June 2006. BE was defined as the presence of specialized intestinal metaplasia (that is, columnar epithelium with goblet cells) in an esophageal biopsy taken from the tubular esophagus by upper gastrointestinal endoscopy, irrespective of the length of involvement. BE patients were prospectively identified from the private and public pathology laboratories servicing metropolitan Brisbane. Of 1714 patients with presumptive BE, we gained permission to contact 1096 (64% response rate). Of these, 614 were ineligible (487 were 'prevalent cases' with a previous BE diagnosis, 86 had only intestinal metaplasia of the gastroesophageal junction, 30 invalid address, 6 too old, 5 other) and 89 were excluded from the study (3 too ill, 5 unable to complete an English language questionnaire, 5 unable to be contacted, 76 failed to return a completed questionnaire). Thus, 393 BE cases (285 nondysplastic, 108 dysplastic) completed the study.

Two separate control groups were recruited for the Study of Digestive Health: 'inflammation controls' (i.e., patients who underwent endoscopy but for whom the histology report

identified only esophageal inflammatory changes consistent with GER and no other pathological or macroscopic changes including no evidence of BE), and population controls. For the purposes of developing risk prediction models to discriminate patients with BE from GERD patients without BE, the present analyses compared the BE cases to the inflammation controls only. In total, 706 of 1354 patients approached as inflammation controls gave permission to be contacted (52% response rate). Of these, 57 refused to participate, 317 were ineligible (304 previous diagnosis of inflammation of the esophagus, 11 invalid address, 1 too old, 1 other), 19 were excluded (6 uncontactable, 1 psychological problems, 1 too ill, 3 unable to complete an English language questionnaire, 8 other), and 313 completed the study.

Both cases and inflammation controls were ineligible for the study if they had a previous diagnosis of BE or cancer. Approval to undertake the Study of Digestive was obtained from the human research ethics committees of the Queensland Institute of Medical Research and all participating hospitals and written informed consent was obtained from all participants.

Candidate predictor variables

Candidate predictor variables were selected *a priori* from the literature and practitioner input, and included: age (years); highest level of education (school only, technical college/diploma, university); body mass index (BMI) 1 year prior to diagnosis (<25, 25-29.9, ≥30 kg/m^2); smoking status (never smoker, ex-smoker, current smoker); cumulative smoking exposure (never smoker, 0-29.9, ≥30 pack-years); smoking duration (never smoker, <15, 15-24, 25-34, ≥35 years); average lifetime alcohol consumption (non-drinker, <1, 1-6, 7-20, ≥21 drinks/week); frequency of use of acid suppressant medications (including proton pump inhibitors and H$_2$-receptor antagonists) in the past 5 years (never, ever); frequency of use of aspirin and other non-steroidal anti-inflammatory drugs (NSAIDs) in the past 5 years (never,

less than weekly, at least weekly); physical activity levels (low, medium, high); average fruit (<2, ≥2 serves/day) and vegetable (<3, ≥3 serves/day) consumption; and number of co-morbidities (categories defined by Charlson et al. (30)). A standardized health and lifestyle questionnaire was used to collect detailed information on these variables for each participant. Most items in the questionnaire showed excellent repeatability after four months (31). Furthermore, we conducted a follow-up interview with the BE cases up to seven years after diagnosis and found similar self-reports of key characteristics (κ ranged from 0.65 to 0.80), suggesting very high reproducibility for these measures. We imputed data for the small proportion of participants with missing values. We compared the model with imputed data with a complete case analysis and found similar model coefficients, but more precise estimates with imputed data.

Validation dataset

The prediction model was externally validated using data from a community-based case-control study of BE conducted in western Washington State, USA (29). BE cases were defined as residents aged 20-80 years newly diagnosed with BE (i.e., specialized intestinal metaplasia in an esophageal biopsy). Of the 208 patients diagnosed with BE, 193 (92.8%) were successfully interviewed. We subsequently excluded 18 cases who were simultaneously diagnosed with esophageal adenocarcinoma (n=2) and/or dysplasia (n=16) from the validation analysis. GERD controls were a random sample of patients (~50%) who underwent endoscopy for reflux symptoms, but who were biopsy-proven negative for BE. Of the 463 patients selected to be GERD controls, 418 (90.8%) were successfully interviewed and were included in the validation analysis.

Statistical analysis

We used basic descriptive statistics to characterize the study populations. For comparisons between BE cases and inflammation controls, we used the χ^2 test for categorical variables and the Student's *t*-test for continuous variables. Statistical computations were performed using SAS software (version 9.2; SAS Institute, Cary, NC), and all tests for statistical significance were two-sided at $\alpha = 0.05$.

Risk model development

We developed separate risk models to predict nondysplastic BE and dysplastic BE in patients with GER symptoms. However, we were unable to externally validate the results for dysplastic BE and report only the nondysplastic BE model here (see supplementary material for the dysplastic BE risk model). As there is a sex difference in the incidence of BE and as some risk factors appear to have different associations with BE in men and women, we identified the important sex-specific predictors for BE and then pooled these in an overall risk model which included a term for sex. The predictors were identified using two phases of stepwise backwards logistic regression. In the first phase, we included in the multivariate model those variables that were statistically significantly associated with BE at the 5% level in univariate analyses and we performed a backwards stepwise regression procedure, whereby those factors losing their significance in the multivariate analysis were dropped. In the second phase, those factors not significant in the univariate analyses were subsequently fitted to the multivariate model to identify those effects detectable only after adjusting for the major risk factors. There was no evidence of multicollinearity between the final list of predictor variables (all variance inflation factors $<< 10$) and no interaction terms were statistically significant (all *p*-values for the type III analysis of effects for interaction terms were >0.10).

Model validation

The accuracy of the model was assessed using tests for discrimination and calibration (32).

We evaluated predictive discrimination using the area under the received operator

characteristic curve (AUC; also known as the c-statistic) and its 95% confidence interval

(95% CI). The AUC can be interpreted as the probability that the model will assign a higher

probability of actually having BE to a randomly chosen patient with BE than to a randomly

chosen patient without BE, or simply the ability of the model to separate cases and

inflammation controls. An AUC of 0.5 indicates that the model has a predictive

discrimination no better than chance, whereas an AUC of 1.0 indicates a perfectly

discriminating model. The second measure calculated was calibration, which compares the

predicted probability with the observed risk. When the average predicted risk within decile

categories matches the proportion observed, the model is well calibrated. We evaluated

calibration using the Hosmer-Lemeshow goodness-of-fit statistic (33), where a high p-value

indicates excellent calibration. Calibration curves were also plotted to illustrate the model's

fit across the range of predicted risk for BE compared with the observed outcome.

RESULTS

Epidemiologic data were available for the derivation analysis from 285 nondysplastic BE

patients (cases) and 313 inflammation controls (Table 1). BE cases were more likely to be

men (64% vs. 47%, *p*<0.001) and were, on average, 5 years older than inflammation controls.

The majority (96%-97%) of cases and inflammation controls reported being Caucasian

(*p*=0.50).

All the potential predictive covariates with their univariate analyses are presented in Table 2.

In the univariate analyses among men (181 cases, 147 inflammation controls), highest level

of education, BMI, tobacco smoking, and frequency of use of acid suppressant medications

were all statistically significantly associated with BE risk. In women (104 cases, 166

inflammation controls), only highest level of education and frequency of use of acid

suppressant medications were significantly associated with BE risk in the univariate analyses.

Alcohol consumption, frequency of use of NSAIDs, physical activity, fruit and vegetable

consumption, and the presence of co-morbidities were not statistically significantly

associated with BE for men or women in the univariate analyses.

The variables retained in the final multivariate risk model included age, sex, smoking status,

BMI, highest level of education and frequency of use of acid suppressant medications (Table

3). There were no statistically significant interactions between sex and the other variables in

the final model.

Model performance

The risk prediction model for nondysplastic BE had good discrimination, with an AUC of

0.70 (95%CI 0.66-0.74) in the development dataset. The discriminatory performance of the

model in the validation dataset was more moderate with an AUC of 0.61 (95%CI 0.56-0.66). Performance of the model was statistically good by the goodness-of-fit test (Hosmer-Lemeshow test, p=0.75) and the calibration curve (Figure 1) shows good agreement between predicted probabilities and actual BE risk across the observed range of risk.

Clinical application using probability thresholds

To assess the potential effects of using the prediction model to guide referral for endoscopic screening for BE, we calculated the proportion of patients that would be referred for endoscopy at different probability thresholds for BE (Table 4). The first row gives the scenario of referring every patient with GER symptoms for endoscopy and therefore identifying all patients with GER symptoms who have BE (that is, sensitivity is 100%). If patients are referred for endoscopy only if their predicted probability is, for example, 50% or more, the proportion of patients referred for endoscopy will be reduced to 46%. At that threshold however, about 41% of BE cases (sensitivity, 59%) will not be referred for endoscopy. As the threshold increases, the number of referrals is reduced; as a consequence, however, the number of patients with BE who will not be referred for endoscopy increases.

DISCUSSION

In this study, we developed and externally validated a clinical risk prediction model for BE based on existing data from a large, population-based study. We used a rigorous statistical approach to determine the most important panel of risk factors to predict the presence of BE in patients with GER symptoms. As recent epidemiological studies have shown that some risk factors, notably obesity, appear to have different associations with BE in men and women, we identified sex-specific predictors and then pooled these in an overall model. The final risk model included terms for age, sex, smoking status, BMI, highest level of education and frequency of use of acid suppressant medications. External validation of the model showed that it performed moderately well in discriminating between patients with nondysplastic BE complicating their GER (cases) and those with no BE (inflammation controls), and the predicted risk correlated well with the observed risk.

As BE is assumed to be an intermediate step in the development of EAC, screening to identify people with BE may be an effective strategy to prevent progression to cancer, at least in theory. Given the high prevalence of GER symptoms in the population and low prevalence of BE in these patients, endoscopic screening for BE in all patients with GER symptoms is not recommended in current international guidelines (13, 14, 34). The American Gastroenterological Association guideline recommends screening in selective populations with multiple risk factors for EAC (age 50 years or older, male sex, white race, chronic GER, hiatal hernia, and elevated BMI) (14). The evidence base underpinning these guidelines is not strong however, and adherence to the guidelines is likely to be incomplete. Data for Australia are limited, however a New Zealand study showed that approximately 50% of indications for endoscopy in 2003 were for heartburn or dyspepsia (i.e., to exclude BE/EAC), significantly higher than in 1997 (35). In healthcare systems throughout the world, there is

an increasing need for evidence based strategies including the need to establish an effective means of risk stratification for endoscopic BE screening among patients with GER symptoms.

Risk prediction can be used in clinical settings to stratify individuals into homogeneous risk groups. Risk prediction models are used in public health to quantify the probability of disease based on a combination of risk factors (36). So far, risk prediction approaches have been used extensively for cardiovascular diseases (16), and more recently, there has been a focus on deriving cancer risk prediction models (17-26). These models can complement clinical assessment, but they also ensure that the decision making process is more uniform across different centers by moving away from using any individual clinician's personal experience (37).

Previous efforts to develop a risk model for BE focused on identifying BE among GER patients using gastrointestinal symptoms (27, 28). While these models performed well (AUCs > 0.70), they have not been externally validated. Our study utilized similar methods to derive a risk prediction model for BE using phenotypic and environmental risk factors, and tested its performance in an external population. The variables included in the model are all important and not necessarily causal correlates of BE and are supported by published findings from our own and other case-control and cohort studies of BE (10, 12, 38-41). Furthermore, to encourage generalizability, we emphasized the use of information on risk factors that can be obtained by practitioners in the office setting during routine healthcare. Importantly, the discriminatory accuracy for our model (AUC=0.61 in the external validation dataset) compares favorably with cancer risk prediction models, such as the Gail (42) model for

Breast cancer risk (0.58) (43), and the LLP (22), Spitz (23), and Bach (21) lung cancer risk models (0.69, 0.69, and 0.62, respectively) (44).

Applying this model to all patients with GER symptoms currently being considered for endoscopy, and using a threshold for making a decision has the potential to reduce the number of unnecessary endoscopies performed to exclude BE. This model makes explicit the proportion of BE cases who would be missed because their predicted risk lies below the threshold. Due to the case-control study design, we were unable to determine precisely the positive and negative predictive values for the model. However, if we assume 5% prevalence of BE among persons with GER symptoms, our best estimate is that 5-17% of persons referred for endoscopy will have BE (depending upon the cut point chosen for referral) and 95-100% of persons not referred will not have BE. Assuming 15% prevalence, 15-41% of persons referred for endoscopy will have BE and 85-100% of people not referred will not have BE. 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 BE model, as the absolute risk of progression to cancer in BE patients is low (8), a threshold whereby fewer investigations are performed at the risk of missing more BE cases (i.e., increased specificity and decreased sensitivity) may be desirable.

Our study had strengths and limitations. The large samples of patients newly diagnosed with BE in the two settings were recruited prospectively, and comparable, consistent and standardized criteria were used for histologic and endoscopic definitions of BE. Both the derivation sample and the validation sample only included patients who were selected for endoscopy; the large (but unknown) proportion of patients with GER symptoms *not* referred

for screening had already been triaged away from endoscopy by clinicians using their own internal algorithms. Presumably, the clinicians had decided that those patients were at such low risk of significant pathology that there was no net benefit from undergoing endoscopic investigation. As such, it is likely that had those low risk patients been included in the two samples, our prediction models would have performed even better. While our modeling assumes that endoscopy is performed in the setting of GER symptoms solely to exclude BE, endoscopy may be undertaken for other indications in this clinical setting. If so, then this would tend to attenuate the predictive value of the models we have derived, since those patients being referred for other indications would presumably be at lower risk of BE than those being referred to confirm the clinical diagnosis.

A limitation of the Australian study was the relatively low rate of participation, raising concerns about possible biased selection of cases and controls. Because BE cases and inflammation controls were sampled from the same population, navigated the same clinical pathways and were recruited using identical methods, it is unlikely that systematically biased selection of one or other group explains our findings. Moreover, BE cases and inflammation controls were not informed about the hypotheses being tested, and so while biased recall of non-reflux exposures is possible, we consider the likelihood that this accounts for our observations as low. Although there were 108 dysplastic BE cases in our development dataset, we were unable to obtain a validation dataset with dysplastic BE cases. The best estimate for the performance of our model for predicting BE with dysplasia was an AUC estimate of 0.87 (95%CI 0.83-0.91) in the development dataset. Recently, central obesity has been found to be more strongly associated with BE than BMI, however measures of central obesity (e.g., waist-to-hip ratio) were not collected for this study. It is likely that adding such measures to the model, rather than BMI, may improve predictive accuracy. Finally, our

sample was predominantly white (~97%) and thus the models may not be applicable to other ethnic groups.

This parsimonious model however could be considered as a starting point for further development, as a number of risk factors were not included and genetic information may also be important in predicting risk of BE. The inclusion of other environmental risk factors and the extension of the model to include biomarkers may go further to improving performance. However, a recent study has shown that breast cancer risk prediction does not improve significantly when genetic information is included in the risk model (45).

In summary, we have derived and externally validated a risk prediction model which estimates the likelihood of undiagnosed BE in patients with GER symptoms being considered for upper gastrointestinal endoscopy. The prediction model has the potential to be a useful tool in the clinical setting for decisions regarding investigation and treatment of patients with GER.

REFERENCES

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127:2893-917.

2. Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al., editors. Cancer Incidence in Five Continents, Vol. IX: IARC Scientific Publications No. 160, Lyon, IARC.; 2007.

3. Pohl H, Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. J Natl Cancer Inst 2005;97:142-6.

4. Cancer Research UK. Oesophageal cancer: UK incidence statistics 2010.   [14 November 2011]; Available from:

http://info.cancerresearchuk.org/cancerstats/types/oesophagus/incidence/.

5. Stavrou EP, McElroy HJ, Baker DF, Smith G, Bishop JF. Adenocarcinoma of the oesophagus: incidence and survival rates in New South Wales, 1972-2005. Med J Aust 2009;191:310-4.

6. O'Shaughnessy JA, Kelloff GJ, Gordon GB, Dannenberg AJ, Hong WK, Fabian CJ, et al. Treatment and prevention of intraepithelial neoplasia: an important target for accelerated new agent development. Clin Cancer Res 2002;8:314-46.

7. Shaheen NJ, Crosby MA, Bozymski EM, Sandler RS. Is there publication bias in the reporting of cancer risk in Barrett's esophagus? Gastroenterology 2000;119:333-8.

8. Hvid-Jensen F, Pedersen L, Drewes AA, Sørensen HT, Funch-Jensen P. Incidence of adenocarcinoma among patients with Barrett's esophagus. N Engl J Med 2011;365:1375-83.

9. Falk GW. Barrett's esophagus. Gastroenterology 2002;122:1569-91.

10. Corley DA, Kubo A, Levin TR, Block G, Habel L, Zhao W, et al. Abdominal obesity and body mass index as risk factors for Barrett's esophagus. Gastroenterology 2007;133:34-41.

11. Edelstein ZR, Farrow DC, Bronner MP, Rosen SN, Vaughan TL. Central adiposity and risk of Barrett's esophagus. Gastroenterology 2007;133:403-11.

12. Smith KJ, O'Brien SM, Green AC, Webb PM, Whiteman DC. Current and past smoking significantly increase risk for Barrett's esophagus. Clin Gastroenterol Hepatol 2009;7:840-8.

13. Guidelines for the diagnosis and management of Barrett's columnar-lined oesophagus. A report of the working party of the British Society of Gastroenterology. [cited 2011 1 November]; Available from: http://www.bsg.org.uk.

14. Spechler SJ, Sharma P, Souza RF, Inadomi JM, Shaheen NJ, Allen JI, et al. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. Gastroenterology 2011;140:1084-91.

15. Sharma P, Sidorenko EI. Are screening and surveillance for Barrett's oesophagus really worthwhile? Gut 2005;54:I27-I32.

16. Grundy SM, Balady GJ, Criqui MH, Fletcher G, Greenland P, Hiratzka LF, et al. Primary prevention of coronary heart disease: guidance from Framingham. Circulation 1998;97:1876-87.

17. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81:1879-86.

18. Colditz GA, Rosner B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. Am J Epidemiol 2000;152:950-64.

19. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 2004;23:1111-30.

20. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst 2006;98:1204-14.

21. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, et al. Variations in lung cancer risk among smokers. J Natl Cancer Inst 2003;95:470-8.

22. Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, et al. The LLP risk model: an individual risk prediction model for lung cancer. Br J Cancer 2008;98:270-6.

23. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, et al. A risk model for prediction of lung cancer. J Natl Cancer Inst 2007;99:715-26.

24. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, et al. Assessing prostate cancer risk: results from the prostate cancer prevention trial. J Natl Cancer Inst 2006;98:529-34.

25. Wu X, Lin J, Grossman HB, Huang M, Gu J, Etzel CJ, et al. Projecting individualized probabilities of developing bladder cancer in white individuals. J Clin Oncol 2007;25:4974-81.

26. Fears TR, Guerry D, Pfeiffer RM, Sagebiel RW, Elder DE, Halpern A, et al. Identifying individuals at high risk of melanoma: a practical predictor of absolute risk. J Clin Oncol 2006;24:3590-6.

27. Gerson LB, Edson R, Lavoir PW, Triadafilopoulos G. Use of a simple symptom questionnaire to predict Barrett's esophagus in patients with symptoms of gastroesophageal reflux. Am J Gastroenterol 2001;96:2005-12.

28. Locke GR, Zinsmeister AR, Talley NJ. Can symptoms predict endoscopic findings in GERD? Gastrointest Endosc 2003;58:661-70.

29. Edelstein ZR, Bronner MP, Rosen SN, Vaughan TL. Risk factors for Barrett's esophagus among patients with gastroesophageal reflux disease: a community clinic-based case-control study. Am J Gastroenterol 2009;104:834-42.

30. Charlson ME, Pompei P, Ales KL, Mackenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373-83.

31. Karatela S, Purdie DM, Green AC, Webb PM, Whiteman DC. Repeatability of self-reported information for population-based studies of cancer. Asian Pac J Cancer Prev 2006;7:303-8.

32. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-38.

33. Hosmer DW, Hosmer T, leCessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 1997;16:965-80.

34. Shaheen N, Ransohoff DF. Gastroesophageal reflux, Barrett esophagus, and esophageal cancer - Scientific review. JAMA 2002;287:1972-81.

35. Schultz M, Davidson A, Donald S, Targonska B, Turnbull A, Weggery S, et al. Gastroenterology service in a teaching hospital in rural New Zealand, 1991-2003. World J Gastroenterol 2009;15:583-90.

36. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.

37. Domchek SM, Eisen A, Calzone K, Stopfer J, Blackwood A, Weber BL. Application of breast cancer risk prediction models in clinical practice. J Clin Oncol 2003;21:593-601.

38. Taylor JB, Rubenstein JH. Meta-analyses of the effect of symptoms of gastroesophageal reflux on the risk of Barrett's esophagus. Am J Gastroenterol 2010;105:1730-7.

39. Thrift AP, Pandeya N, Smith KJ, Green AC, Webb PM, Whiteman DC. The use of nonsteroidal anti-inflammatory drugs and the risk of Barrett's oesophagus. Aliment Pharmacol Ther 2011;34:1235-44.

40. Thrift AP, Pandeya N, Smith KJ, Mallitt KA, Green AC, Webb PM, et al. Lifetime alcohol consumption and risk of Barrett's esophagus. Am J Gastroenterol 2011;106:1220-30.

41. Thompson OM, Beresford SAA, Kirk EA, Vaughan TL. Vegetable and fruit intakes and risk of Barrett's esophagus in men and women. Am J Clin Nutr 2009;89:890-6.

42. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999;91:1541-8.

43. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 2001;93:358-66.

44. D'Amelio AM, Jr., Cassidy A, Asomaning K, Raji OY, Duffy SW, Field JK, et al. Comparison of discriminatory power and accuracy of three lung cancer risk models. Br J Cancer 2010;103:423-9.

45. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, et al. Performance of common genetic variants in breast-cancer risk models. N Engl J Med 2010;362:986-93.

Table 1. Socio-demographic characteristics of inflammation controls and nondysplastic BE cases

Variables	Inflammation controls (n=313)	Nondysplastic BE cases (n=285)	p-value[a]
Mean age (SD[b]), y	53.5 (12.7)	58.2 (11.9)	<.001
Sex, n (%)			
Men	147 (47.0)	181 (63.5)	
Women	166 (53.0)	104 (36.5)	<.001
Ethnicity, n (%)			
Caucasian	304 (97.1)	274 (96.1)	
Other	9 (2.9)	11 (3.9)	.50

BE, Barrett's esophagus.
[a]p-value from the two-sided χ^2 test (categorical variables) and t-test (continuous variables).
[b]SD, standard deviation.

Table 2. Univariate analyses of nondysplastic Barrett's esophagus risk factors by sex

Risk factor	Inflammation controls (n=147) n (%)	Nondysplastic BE cases (n=181) n (%)	Crude OR (95% CI)	p^a	Inflammation controls (n=166) n (%)	Nondysplastic BE cases (n=104) n (%)	Crude OR (95% CI)	p^a
	Men				**Women**			
Highest level of education								
School only	37 (25.2)	68 (37.6)	2.59 (1.37-4.88)		79 (47.6)	70 (67.3)	2.95 (1.31-6.65)	
Technical college/Diploma	72 (49.0)	86 (47.5)	1.68 (0.94-3.02)		57 (34.3)	25 (24.0)	1.46 (0.61-3.53)	
University	38 (25.8)	27 (14.9)	1.00	**0.01**	30 (18.1)	9 (8.7)	1.00	**0.006**
Body mass index last year (kg/m^2)								
<25	45 (30.6)	52 (28.7)	1.00		76 (45.8)	36 (34.6)	1.00	
25-29.9	76 (51.7)	74 (40.9)	0.84 (0.51-1.41)		49 (29.5)	38 (36.5)	1.64 (0.92-2.93)	
≥30	26 (17.7)	55 (30.4)	1.83 (0.99-3.38)	**0.03**	41 (24.7)	30 (28.9)	1.55 (0.84-2.86)	0.19
Smoking status								
Never smoker	69 (46.9)	54 (29.8)	1.00		82 (49.4)	44 (42.3)	1.00	
Ex-smoker	64 (43.5)	95 (52.5)	1.90 (1.18-3.06)		56 (33.7)	42 (40.4)	1.40 (0.81-2.40)	
Current smoker	14 (9.5)	32 (17.7)	2.92 (1.42-6.01)	**0.004**	28 (16.9)	18 (17.3)	1.20 (0.60-2.40)	0.48
Average lifetime alcohol consumption (standard drinks/week)^b								
Non-drinker	10 (6.8)	14 (7.7)	1.00		22 (13.3)	23 (22.1)	1.00	
<1	3 (2.0)	9 (5.0)	2.14 (0.46-9.98)		21 (12.6)	13 (12.5)	0.59 (0.24-1.47)	
1-6	43 (29.3)	45 (24.9)	0.75 (0.30-1.86)		67 (40.4)	49 (47.1)	0.70 (0.35-1.40)	
7-20	51 (34.7)	56 (30.9)	0.78 (0.32-1.92)		48 (28.9)	17 (16.4)	0.34 (0.15-0.76)	
≥21	40 (27.2)	57 (31.5)	1.02 (0.41-2.52)	0.51	8 (4.8)	2 (1.9)	0.24 (0.05-1.25)	0.06
Frequency of use of acid suppressant medications (H$_2$Rs or PPIs) in the past 5 years								
Never	102 (69.4)	88 (48.6)	1.00		108 (65.1)	50 (48.1)	1.00	
Ever	45 (30.6)	93 (51.4)	2.40 (1.52-3.78)	**<.001**	58 (34.9)	54 (51.9)	2.01 (1.22-3.32)	**0.006**
Frequency of use of NSAIDS in the past 5 years								
Never	28 (19.0)	48 (26.5)	1.00		40 (24.1)	25 (24.0)	1.00	
Less than weekly	87 (59.2)	93 (51.4)	0.62 (0.36-1.08)		88 (53.0)	42 (40.4)	0.76 (0.41-1.42)	
At least weekly	32 (21.8)	40 (22.1)	0.73 (0.38-1.41)	0.24	38 (22.9)	37 (35.6)	1.56 (0.79-3.06)	0.06

Risk factor	Men				Women			
	Inflammation controls (n=147) n (%)	Nondysplastic BE cases (n=181) n (%)	Crude OR (95% CI)	p^a	Inflammation controls (n=166) n (%)	Nondysplastic BE cases (n=104) n (%)	Crude OR (95% CI)	p^a
Physical activity index								
Low	35 (23.8)	44 (24.3)	1.13 (0.64-2.01)		54 (32.5)	31 (29.8)	0.96 (0.52-1.78)	
Medium	57 (38.8)	76 (42.0)	1.20 (0.73-1.98)		55 (33.1)	39 (37.5)	1.19 (0.66-2.15)	
High	55 (37.4)	61 (33.7)	1.00	0.77	57 (34.3)	34 (32.7)	1.00	0.76
Average fruit consumption								
<2 serves/day	67 (45.6)	84 (46.4)	1.00		69 (41.6)	39 (37.5)	1.00	
≥2 serves/day	80 (54.4)	97 (53.6)	0.97 (0.63-1.50)	0.88	97 (58.4)	65 (62.5)	1.19 (0.72-1.96)	0.51
Average vegetable consumption								
<3 serves/day	83 (56.5)	114 (63.0)	1.00		73 (44.0)	51 (49.0)	1.00	
≥3 serves/day	64 (43.5)	67 (37.0)	0.76 (0.49-1.19)	0.23	93 (56.0)	53 (51.0)	0.82 (0.50-1.33)	0.42
Number of co-morbidities								
None	33 (22.4)	33 (18.2)	1.00		41 (24.7)	22 (21.1)	1.00	
≥1	114 (77.6)	148 (81.8)	1.30 (0.76-2.23)	0.34	125 (75.3)	82 (78.9)	1.22 (0.68-2.20)	0.50

BE, Barrett's esophagus; CI, confidence interval; H₂R, H₂-receptor antagonist; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; PPI, proton pump inhibitor.

[a]p-value for χ^2 test for heterogeneity for comparing group of cases to the inflammation controls for the distribution of each categorical variable.

[b]One standard drink is equivalent to 10g of alcohol.

Table 3. Final multivariate logistic model for nondysplastic Barrett's esophagus

Risk factor	Adjusted[a] OR (95% CI)
Age	
Per 5 years	1.14 (1.06-1.23)
Sex	
Women	1.00 (Ref)
Men	2.17 (1.50-3.14)
Smoking status	
Never smoker	1.00 (Ref)
Ex-smoker	1.41 (0.96-2.06)
Current smoker	1.93 (1.15-3.24)
Body mass index last year (kg/m^2)	
< 25	1.00 (Ref)
25-29.9	0.96 (0.64-1.44)
≥ 30	1.41 (0.90-2.22)
Education	
University	1.00 (Ref)
Technical College / Diploma	1.29 (0.77-2.15)
School only	2.08 (1.23-3.50)
Frequency of use of acid suppressant medications (H_2Rs or PPIs) in the past 5 years	
Never	1.00 (Ref)
Ever	2.07 (1.46-2.93)
AUC (95%CI)	
Development dataset	0.70 (0.66-0.74)
Validation dataset	0.61 (0.56-0.66)

AUC, area under the receiver operator characteristic curve; CI, confidence interval; H_2R, H_2-receptor antagonist; OR, odds ratio; PPI, proton pump inhibitor.
[a]Odds ratios were adjusted for all other terms in the table.

Table 4. Performance of various probability thresholds for referring patients with GER symptoms for endoscopy for Barrett's esophagus
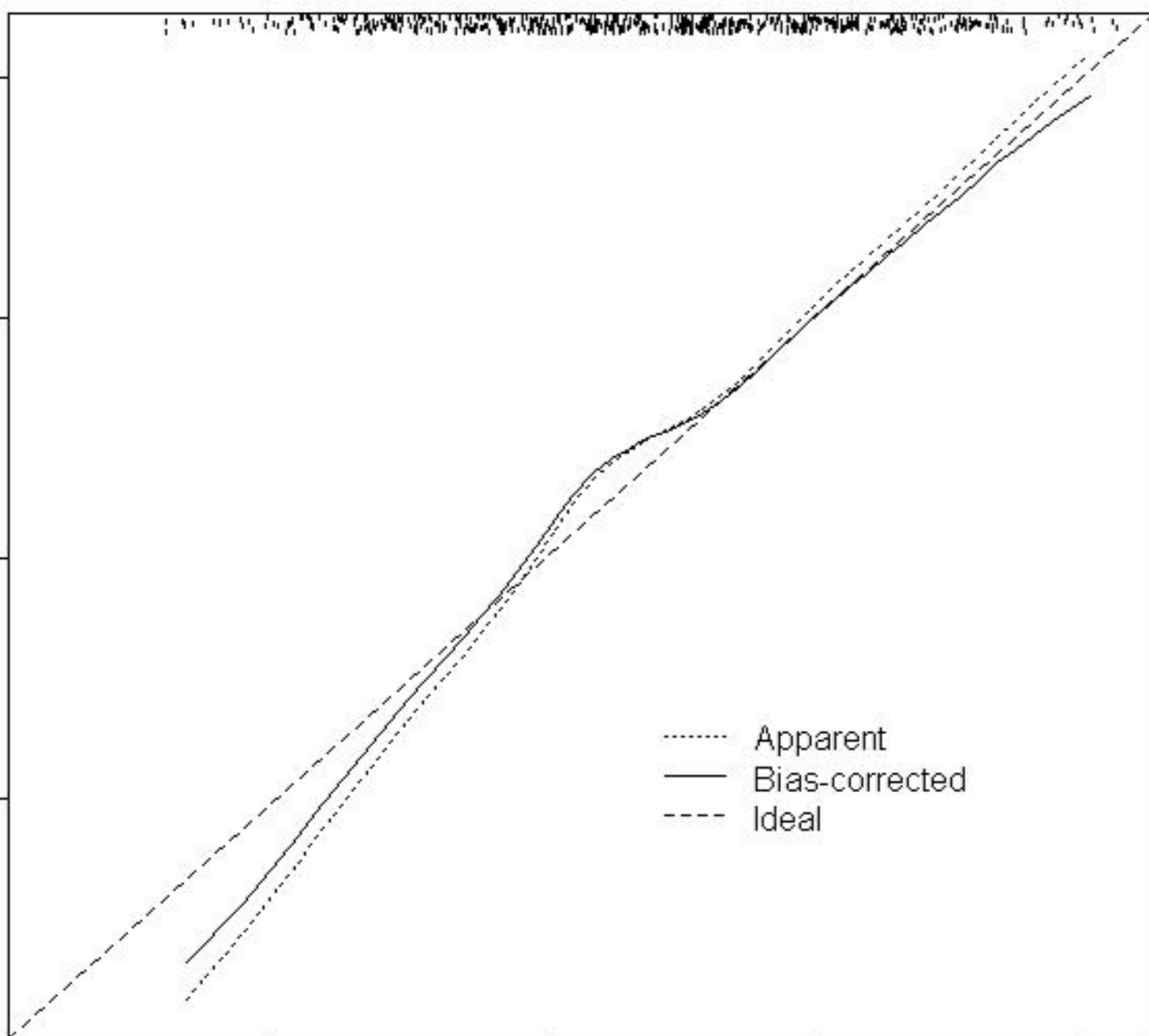
Probability threshold at which a patient with GER is referred for endoscopy	Sensitivity (%)	Specificity (%)	Patients undergoing endoscopy (%)
All	100	0	100
≥0.1	100	0.6	100
≥0.2	97.5	8.3	94.5
≥0.3	91.6	28.4	81.1
≥0.4	75.1	48.9	62.5
≥0.5	59.3	66.5	45.8
≥0.6	38.9	80.5	28.8
≥0.7	17.2	92.7	12.0
≥0.8	3.9	99.0	2.3
≥0.9	0	100	0

GER, gastroesophageal reflux.

Figure 1 Logistic calibration curve for validation of model to predict nondysplastic Barrett's esophagus among men and women presenting with gastroesophageal reflux symptoms.