



Published in final edited form as:

Dig Dis Sci. 2019 December ; 64(12): 3463–3470. doi:10.1007/s10620-019-05707-2.

Use of the Electronic Health Record to Target Patients for Non-endoscopic Barrett’s Esophagus Screening

Brittany L. Baldwin-Hunter¹, Rita M. Knotts², Samantha D. Leeds¹, Joel H. Rubenstein^{3,4}, Charles J. Lightdale¹, Julian A. Abrams¹

¹Division of Digestive and Liver Diseases, Columbia University Medical Center, 630 W 168th Street, P&S 3-401, New York, NY 10032, USA

²Division of Gastroenterology and Hepatology, New York University School of Medicine, 240 East 38th Street, 23rd Floor, New York, NY 10016, USA

³Veterans Affairs Center for Clinical Management Research, Ann Arbor, MI, USA

⁴Barrett’s Esophagus Program, Division of Gastroenterology, University of Michigan, Taubman Center, Floor 3 Reception D, 1500 E Medical Center Dr SPC 2435, Ann Arbor, MI 48109-2435, USA

Abstract

Background—Clinical prediction models targeting patients for Barrett’s esophagus (BE) screening include data obtained by interview, questionnaire, and body measurements. A tool based on electronic health records (EHR) data could reduce cost and enhance usability, particularly if combined with non-endoscopic BE screening methods.

Aims—To determine whether EHR-based data can identify BE patients.

Methods—We performed a retrospective review of patients ages 50–75 who underwent a first-time esophagogastroduodenoscopy. Data extracted from the EHR included demographics and BE risk factors. Endoscopy and pathology reports were reviewed for histologically confirmed BE. Screening criteria modified from clinical guidelines were assessed for association with BE. Subsequently, a score based on multivariate logistic regression was developed and assessed for its ability to identify BE subjects.

Results—A total of 2931 patients were assessed, and BE was found in 1.9%. Subjects who met screening criteria were more likely to have BE (3.3% vs. 1.1%, $p = 0.001$), and the criteria predicted BE with an AUROC of 0.65 (95% CI 0.59–0.71). A score based on logistic regression modeling included gastroesophageal reflux disease, sex, body mass index, and ever-smoker status and identified BE subjects with an AUROC of 0.71 (95% CI 0.64–0.77). Both prediction tools produced higher AUROCs in women than in men.

*Julian A. Abrams ja660@cumc.columbia.edu.

Conflict of interest Julian Abrams has consulted for Medtronic.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10620-019-05707-2>) contains supplementary material, which is available to authorized users.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Conclusions—EHR-based BE risk prediction tools identify BE patients with fair accuracy. While these tools may improve the efficiency of patient targeting for BE screening in the primary care setting, challenges remain to identify high-risk patients for non-invasive BE screening in clinical practice.

Keywords

Esophageal adenocarcinoma; Risk; Prediction; Model; Criteria

Introduction

Esophageal adenocarcinoma (EAC) has rapidly increased in incidence over the past several decades and continues to carry a poor prognosis [1, 2]. Low survival rates are due in part to late detection, with most patients demonstrating advanced disease at diagnosis [3]. Barrett's esophagus (BE) is a known precursor of EAC, and endoscopic surveillance of BE may potentially reduce EAC mortality through early cancer detection as well as endoscopic treatment of precancerous lesions [4–7]. Unfortunately, approximately 90% of EAC patients do not receive a prior BE diagnosis [8, 9]. Thus, the opportunity may be missed to identify and intervene in at-risk patients [5, 10, 11].

Widespread screening of the general population for BE by upper endoscopy is impractical and expensive. Recently, there has been considerable interest in developing non-endoscopic methods to diagnose BE and would ideally be employed in the primary care setting [12–16]. In order to limit the false positive rate of these tests and minimize the number of unnecessary confirmatory endoscopies, appropriate selection of patients at higher risk for EAC is critical. Also, to achieve optimal uptake in the primary care setting, identification of appropriate patients for non-endoscopic BE screening will require minimal time and effort on the part of physicians and staff.

Current guidelines from gastroenterology societies suggest that screening for BE should be considered in individuals with multiple EAC risk factors, including gastroesophageal reflux disease (GERD), age > 50 years, white race, male sex, central obesity, cigarette smoking, and a family history of EAC [17, 18]. BE risk prediction models based on many of these factors have been developed [19–25]. However, these models require data collection by the provider in the form of a patient interview, questionnaire, and/or measurement of waist and hip circumference, all of which may prove impractical in the clinical setting.

The electronic health record (EHR) contains readily available data that could allow for identification of patients at higher risk for EAC without additional provider input. Compared to existing BE prediction models, use of an EHR-based tool to identify screen-eligible patients could reduce cost and increase ease and speed of administration, particularly if combined with non-endoscopic BE screening methods in the primary care setting. The aim of this study was to determine whether patient characteristics that can be readily extracted from the EHR can predict which individuals are most likely to have BE within a diverse patient population who had undergone a first upper endoscopy.

Methods

Study Population

We performed a retrospective chart review of patients ages 50–75 who underwent a first-time esophagogastroduodenoscopy (EGD) at Columbia University Medical Center (CUMC) between January 2015 and December 2017. Patients were excluded if the patient medical history was not contained within our EHR system (Allscripts, Chicago, Illinois) or if the patient had a history of a prior EGD. History of a prior EGD was determined by review of the EHR for an antecedent EGD report, in addition to review of the current EGD indication for the suggestion of a prior endoscopy (e.g., “Follow up of Barrett’s esophagus”). The Institutional Review Board of Columbia University Medical Center approved this study.

Chart Abstraction

Manual review of the EHR was performed for all patients who had undergone a first EGD. The following chart sections were reviewed: active problem list, past medical history, family history, social history, medications, and vital signs. Individual provider notes were not reviewed, as we felt that the time and effort required would not facilitate efficient targeting of patients for BE screening in clinical practice. We extracted data on the following variables: age, sex, race, ethnicity, body mass index (BMI), smoking history, family history, history of a diagnosis of GERD, history of a typical GERD symptom (heartburn or regurgitation), use of a proton-pump inhibitor (PPI) medication within 1 year prior to EGD, and use of an H2-receptor antagonist (H2RA) medication within 1 year prior to EGD.

The indication for each EGD was recorded and categorized as yes or no for GERD. If any of the listed indications suggested that the patient had a history of GERD, then indication was categorized as yes for GERD. All EGD and pathology reports were manually reviewed. The presence of BE was defined as endoscopic evidence of BE based on the endoscopy report, combined with the presence of intestinal metaplasia on esophageal biopsies. “Irregular Z-lines” were not counted as endoscopic evidence of BE. For purposes of these analyses, cases in which EAC or gastroesophageal junction (GEJ) adenocarcinoma was found were counted as BE.

BE Screening Criteria Development and Analysis (Method 1)

In order to generate an EHR-based tool to predict BE, two approaches were taken. The first was the development of an a priori set of criteria, modified from clinical guide-lines, to identify patients likely to have BE. The purpose of this approach was to create a simple, highly efficient checklist which could be used to rapidly target patients for BE screening. The second approach was to use logistic regression modeling to identify EHR-based factors which predict BE. This information could then be used to generate a risk score based on the association of individual factors with BE.

Employing the first approach, patients were identified as hypothetical candidates for BE screening (“eligible” or “ineligible”) according to prespecified criteria, hereafter referred to as “BE screening criteria.” Given the increased prevalence of BE and EAC in men, BE screening criteria were made more permissive for men compared to women [19, 26, 27]. The

risk factors assessed were: GERD, ever-smoker, BMI $\geq 30\text{kg/m}^2$, and first-degree relative with esophageal or gastric cancer. For BMI, a cutoff of 30 kg/m^2 was chosen given that this represents the frequently used definition of obesity and is also associated with an increased risk of BE and EAC [28]. For family history, we chose to broaden the definition to include gastric cancer to capture GEJ cancers that may have been classified in the medical record as gastric cancer. Men were considered screen eligible if any of the risk factors were present. Women were considered screen eligible if GERD was present as well as at least one additional risk factor. The decision to make GERD a mandatory criterion for women was based largely on the fact that GERD is the strongest known modifiable risk factor for BE and EAC.

GERD was classified as present if any of the following were noted in the EHR: GERD diagnosis listed under active problem list or past medical history; typical GERD symptom (heartburn or regurgitation) listed under active problem list or past medical history; PPI listed as an active medication within 1 year prior to EGD; or H2RA listed as an active medication within 1 year prior to EGD.

Categorical variables were compared between groups using Fisher's exact tests, and continuous variables were compared using *t* tests or rank sum tests, where appropriate. The ability of the prespecified BE screening criteria to predict which patients would have BE on upper endoscopy was assessed in the entire study population through calculation of sensitivity, specificity, positive and negative likelihood ratios, and area under the receiver operating characteristic curve (AUROC). Analyses were then repeated stratified by sex. In secondary analyses, the above was repeated but with subjects categorized based on the presence of "high risk" BE lesions, defined as length $\geq 3\text{cm}$ and/or low-grade dysplasia or worse on biopsies.

BE Risk Prediction Model Development and Analysis (Method 2)

For the second approach, we developed an EHR-based BE risk prediction score using multivariate logistic regression analyses. The candidate predictor variables proposed for the model were established risk factors associated with BE and EAC: age, sex, race, BMI, smoking history (ever vs. never), family history, and GERD [19, 26, 27, 29–35]. For GERD, we evaluated combinations of the various EHR-based methods to "define" GERD (GERD diagnosis, GERD symptom, PPI use, H2RA use) to determine whether a particular definition may better identify patients with BE. For each of these combinations, we calculated odds ratios and AUROCs for the association with BE. We performed multivariable logistic regression modeling, including in a full model all variables associated with BE with a *p* value < 0.10 . Smoking history was categorized as ever/never, and BMI was evaluated as a continuous variable. We subsequently created a reduced model, removing variables in a stepwise fashion with highest *p* value that was > 0.15 . We then forced the indication for EGD into the selected multivariable model to assess whether this impacted the results. These analyses were performed for the entire study population and then repeated stratified by sex. For the final models, the optimal cut-points were determined using the Youden's index method [36]. The Youden's index is calculated for each possible cut point as (sensitivity + specificity – 1) and is defined graphically as the vertical distance between the ROC curve

and the diagonal chance line. AUROCs were calculated for the models, and tenfold cross-validation was performed. As described above in Method 1, secondary analyses were performed to assess a risk prediction model for the identification of “high risk” BE lesions. Statistical significance was defined as $p < 0.05$. All analyses were performed using Stata 14.0 (StataCorp, College Station, TX).

Results

Patient Characteristics

A total of 2931 patients underwent a first EGD between 2015 and 2017 and were included in the study analyses. Patient characteristics are shown in Table 1. The majority of patients were female (55.3%), mean BMI was 28.8 (SD 7.5), 36.4% of patients were ever-smokers, and 23.7% had a GERD-related indication for EGD. The prevalence of histologically confirmed BE was 1.9%. The median length of BE was 2.0cm (IQR 1.0–3.0 cm). Forty-three patients (75.4%) had no dysplasia, 5 had indefinite dysplasia (8.8%), 1 had low-grade dysplasia (1.8%), 2 had high-grade dysplasia (3.5%), and 6 had adenocarcinoma (10.5%). Twenty patients had “high risk” BE (length ≥ 3 cm and/or low-grade dysplasia or worse). Patients with BE had a higher mean BMI (30.9 vs. 28.7, $p = 0.04$) and were more likely to be an ever-smoker (54.4 vs. 36.2%, $p = 0.02$), to have a GERD diagnosis (42.1 vs. 23.3%, $p = 0.002$), and to have used a PPI within one year prior to EGD (38.6 vs. 23.0%, $p = 0.01$). No significant difference was seen in the prevalence of GERD symptoms, H2RA use, or family history of esophageal or gastric cancer in patients with and without BE.

EHR-Based BE Screening Criteria

Using the prespecified criteria, 44% of patients would have been eligible for BE screening (68.5% of men, 31.5% of women). Screen-eligible patients had a significantly higher prevalence of BE (3.3 vs. 1.1%, $p = 0.001$). Overall, BE screening criteria discriminated between patients with and without BE with 73.7% sensitivity and 56.7% specificity and produced an AUROC of 0.65 (95% CI 0.59–0.71) and positive and negative likelihood ratios of 1.70 and 0.46, respectively (Table 2). When stratified by sex, the criteria were more sensitive (87.5%) and less specific (56.7%) in men and produced a lower AUROC of 0.60. Among women, the criteria were less sensitive (32.6%) but more specific (76.0%) and produced a higher AUROC of 0.66 (Fig.1). In secondary analyses, the prespecified criteria identified patients with “high risk” BE lesions with 90.0% sensitivity, 56.4% specificity, and an AUROC of 0.73 (95% CI 0.66–0.80).

Development of EHR-Based BE Risk Prediction Model

In order to determine the best method to parameterize GERD from the EHR, various combinations of GERD-related variables were assessed as predictors of BE (Supplementary Table 1). Addition of GERD symptoms to GERD diagnosis did not further increase the AUROC. Interestingly, all of the potential GERD definitions had numerically higher AUROCs in women as compared to men. The definition of GERD as (GERD diagnosis or PPI use or H2RA use) produced the highest AUROC (0.62, 95% CI 0.56–0.69), and this definition was used for subsequent analyses to define GERD.

The final multivariable model for association with BE included sex, smoking history, BMI, and GERD and produced an AUROC of 0.71 (95% CI 0.64–0.77) (Fig.2 and Table 3). In cross-validation, the AUROC for the model was 0.66 (95% CI 0.59–0.73). The optimal cut-point from this model was associated with 79.2% sensitivity and 58.0% specificity. Removing sex from the model and then stratifying by sex, the AUROCs were again numerically higher in women than in men. In women, the AUROC was 0.77 (95% CI 0.69–0.86; cross-validation AUROC 0.75), while in men the AUROC was 0.61 (95% CI 0.51–0.71, cross-validation AUROC 0.55).

In secondary analyses, a risk prediction model was developed and analyzed for the identification of “high risk” BE lesions. The final multivariable model included sex, smoking history, and GERD. This model produced an AUROC of 0.76 (95% CI 0.68–0.85) and was 0.70 (95% CI 0.62–0.79) in cross-validation. The optimal cut-point from this model for “high risk” BE was associated with 85.0% sensitivity and 63.2% specificity.

Discussion

In the current study, we demonstrated that EAC risk factors identified from patient electronic health records could be used to identify patients who had BE on upper endoscopy. The first approach was an a priori set of BE screening criteria (the “checklist” approach) that discriminated between BE and non-BE patients with fair accuracy. With the second approach, we developed a risk score based on a logistic regression modeling that included history of GERD, BMI, and ever-smoker status, and this score was similar to the a priori BE screening criteria in predicting the presence of BE. Interestingly, both approaches were more accurate in women than in men.

Several prior studies have put forth BE risk prediction models [20–24]. In general, these models included data elements that required patient or provider input. Frequency, duration, and severity of GERD symptoms were generally assessed by questionnaire. Waist-to-hip ratio requires body measurements that are not routinely performed in clinical practice. Pack years of smoking and highest level of education are not routinely and accurately recorded in the medical records. Few models have been put forth to identify patients at highest risk for EAC, and most of these models also require information that is not readily obtained from the EHR [37–39]. Recently, a Swedish case control study has developed a simplified EAC risk prediction model, which could theoretically be applied to the EHR [39]. The variables included in this model were BMI, smoking status, and the presence of either GERD symptoms or acid suppressant medication use. In our analyses, however, EHR-derived GERD symptoms were a poor predictor of any BE and of high-risk BE. The BE prediction tools proposed in the current study were generated solely using EHR-derived data and performed comparably to existing clinical prediction models.

BE screening practices may evolve dramatically in the near future as the development of minimally invasive BE screening methods continues to advance. These methods may offer a low-cost, low-risk alternative to endoscopic BE screening and can potentially be administered in the primary care setting. The best studied of these methods is the Cytosponge, a cell sampling device that, in combination with immunohistochemical staining

for trefoil factor 3, has been shown to diagnose BE with 80% sensitivity and 92% specificity [16]. Additional minimally invasive tests for BE under investigation include use of tethered sampling devices to assess methylated DNA markers, breath testing, and analysis of the salivary microbiome [13–15]. As investigations continue of these novel BE screening methods for use in the primary care setting, questions remain as to which patients should be offered these tests, and how primary care providers, with limited resources and competing healthcare tasks, can identify these patients efficiently. EHR-derived data represent an inexpensive and efficient means to identify these patients. Additionally, many current EHR systems have the capacity to extract relevant data to generate electronic alerts for providers that would identify patients who are candidates for screening. However, it is unclear whether a non-endoscopic screening strategy will ultimately be superior to an endoscopic screening strategy to identify patients at high risk for EAC. While likely cheaper and more efficient, a non-endoscopic testing strategy would be impacted by the sensitivity and specificity of an EHR-based approach to identify patients at risk for BE, combined with the test characteristics of a non-endoscopic test.

Unexpectedly, the BE prediction criteria and model-based score were more accurate at predicting BE in women than in men. Within our study population, GERD, smoking, and obesity were all more strongly associated with BE in women than in men (data not shown). The current study evaluated patients presenting for first-time upper endoscopy, and thus, referral bias may explain these findings. For instance, compared to men who undergo upper endoscopy, women may have more frequent or long-standing reflux. Providers may also have a higher threshold to screen for BE in women due to the lower prevalence of BE and EAC in this population. There also may be differences between men and women in terms of GERD symptom reporting.

The present study has several strengths. We analyzed a large sample of patients with and without BE. The study population was diverse with over 50% women, and fewer than 55% white. We used a strict definition of BE that required both endoscopic evidence and histologic confirmation of BE, minimizing any disease misclassification. From a cancer prevention standpoint, the identification of patients at highest risk for EAC would be most useful. While the number of “high risk” BE patients was somewhat small, the EHR-based approaches did identify these individuals with relatively good accuracy.

There were also certain limitations. Selection bias may have been introduced by including patients undergoing first-time upper endoscopy, who may not be representative of patients in the primary care setting. EHR-based data do not capture well GERD severity and duration or quantitative smoking history, and this information may improve identification of BE patients. Similarly, family history of BE or EAC may not be reliably recorded in the EHR.

In conclusion, risk factors for EAC extracted from the EHR identify patients with BE with fair accuracy. While these data do not identify new risk factors for BE, they do suggest that appropriate patient selection for BE screening may be accomplished simply through review of the EHR. However, future studies are required to determine whether EHR-based algorithms can be improved to better identify patients with BE. Particularly, if coupled with non-endoscopic BE screening, the use of EHR-based data represents an inexpensive and

efficient mechanism to identify patients at risk for EAC and can be translated to the primary care setting. Additional studies are warranted to develop automated, EHR-based identification of patients at high risk for EAC who may benefit from screening.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Abrams JA, Sharaiha RZ, Gonsalves L, Lightdale CJ, Neugut AI. Dating the rise of esophageal adenocarcinoma: analysis of connecticut tumor registry data, 1940–2007. *Cancer Epidemiol Biomark Prev.* 2011;20:183–186. 10.1158/1055-9965.EPI-10-0802.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018;68:7–30. 10.3322/caac.21442. [PubMed: 29313949]
- Njei B, McCarty TR, Birk JW. Trends in esophageal cancer survival in United States adults from 1973 to 2009: a SEER database analysis. *J Gastroenterol Hepatol.* 2016;31:1141–1146. 10.1111/jgh.13289 [PubMed: 26749521]
- Cooper GS, Kou TD, Chak A. Receipt of previous diagnoses and endoscopy and outcome from esophageal adenocarcinoma: a population-based study with temporal trends. *Am J Gastroenterol.* 2009;104:1356–1362. 10.1038/ajg.2009.159. [PubMed: 19491849]
- El-Serag HB, Naik AD, Duan Z, et al. Surveillance endoscopy is associated with improved outcomes of esophageal adenocarcinoma detected in patients with Barrett's esophagus. *Gut.* 2016;65:1252–1260. 10.1136/gutjnl-2014-308865. [PubMed: 26311716]
- Qiao Y, Hyder A, Bae SJ, et al. Surveillance in patients with Barrett's esophagus for early detection of esophageal adenocarcinoma: a systematic review and meta-analysis. *Clin Transl Gastroenterol.* 2015;6:e131 10.1038/ctg.2015.58.
- Wang KK, Sampliner RE. Practice parameters Committee of the American College of G. updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *Am J Gastroenterol.* 2008;103:788–797. 10.1111/j.1572-0241.2008.01835.x. [PubMed: 18341497]
- Tramontano AC, Sheehan DF, Yeh JM, et al. The impact of a prior diagnosis of Barrett's esophagus on esophageal adenocarcinoma survival. *Am J Gastroenterol.* 2017;112:1256–1264. 10.1038/ajg.2017.82. [PubMed: 28374815]
- Verbeek RE, Leenders M, Ten Kate FJ, et al. Surveillance of Barrett's esophagus and mortality from esophageal adenocarcinoma: a population-based cohort study. *Am J Gastroenterol.* 2014;109:1215–1222. 10.1038/ajg.2014.156. [PubMed: 24980881]
- Phoa KN, van Vilsteren FG, Weusten BL, et al. Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia: a randomized clinical trial. *JAMA.* 2014;311:1209–1217. 10.1001/jama.2014.2511. [PubMed: 24668102]
- Shaheen NJ, Sharma P, Overholt BF, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. *N Engl J Med.* 2009;360:2277–2288. 10.1056/NEJMoa0808145. [PubMed: 19474425]
- American Gastroenterological A, Spechler SJ, Sharma P, Souza RF, Inadomi JM, Shaheen NJ. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. *Gastroenterology.* 2011;140:1084–1091. 10.1053/j.gastro.2011.01.030. [PubMed: 21376940]
- Chan DK, Zakko L, Visrodia KH, et al. Breath testing for Barrett's esophagus using exhaled volatile organic compound profiling with an electronic nose device. *Gastroenterology.* 2017;152:24–26. 10.1053/j.gastro.2016.11.001. [PubMed: 27825962]
- di Pietro M, Chan D, Fitzgerald RC, Wang KK. Screening for Barrett's esophagus. *Gastroenterology.* 2015;148:912–923. 10.1053/j.gastro.2015.02.012. [PubMed: 25701083]
- Iyer PG, Taylor WR, Johnson ML, et al. Highly discriminant methylated DNA markers for the non-endoscopic detection of Barrett's esophagus. *Am J Gastroenterol.* 2018;113:1156–1166. 10.1038/s41395-018-0107-7. [PubMed: 29891853]

16. Shaheen NJ, Falk GW, Iyer PG, Gerson LB. American College of G. ACG clinical guideline: diagnosis and management of Barrett's esophagus. *Am J Gastroenterol.* 2016;111:30–50. 10.1038/ajg.2015.322. [PubMed: 26526079]
17. Ross-Innes CS, DeBiram-Beecham I, O'Donovan M, et al. Evaluation of a minimally invasive cell sampling device coupled with assessment of trefoil factor 3 expression for diagnosing Barrett's esophagus: a multi-center case-control study. *PLoS Med.* 2015;12:e1001780 10.1371/journal.pmed.1001780. [PubMed: 25634542]
18. Snider EJ, Compres G, Freedberg DE, et al. Barrett's esophagus is associated with a distinct oral microbiome. *Clin Transl Gastroenterol.* 2018;9:135 10.1038/s41424-018-0005-8. [PubMed: 29491399]
19. 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;104:834–842. 10.1038/ajg.2009.137. [PubMed: 19319131]
20. Gerson LB, Edson R, Lavori 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;96:2005–2012. 10.1111/j.1572-0241.2001.03933.x. [PubMed: 11467625]
21. Ireland CJ, Fielder AL, Thompson SK, Laws TA, Watson DI, Esterman A. Development of a risk prediction model for Barrett's esophagus in an Australian population. *Dis Esophagus.* 2017;30:1–8. 10.1093/dote/dox033.
22. Liu X, Wong A, Kadri SR, et al. Gastro-esophageal reflux disease symptoms and demographic factors as a pre-screening tool for Barrett's esophagus. *PLoS ONE.* 2014;9:e94163 10.1371/journal.pone.0094163. [PubMed: 24736597]
23. Locke GR, Zinsmeister AR, Talley NJ. Can symptoms predict endoscopic findings in GERD? *Gastrointest Endosc.* 2003;58:661–670. [PubMed: 14595298]
24. Rubenstein JH, Morgenstern H, Appelman H, et al. Prediction of Barrett's esophagus among men. *Am J Gastroenterol.* 2013;108:353–362. 10.1038/ajg.2012.446. [PubMed: 23318485]
25. Thrift AP, Kendall BJ, Pandeya N, Vaughan TL, Whiteman DC, Study of Digestive H. A clinical risk prediction model for Barrett esophagus. *Cancer Prev Res (Phila).* 2012;5:1115–1123. 10.1158/1940-6207.capr-12-0010. [PubMed: 22787114]
26. Eloubeidi MA, Provenzale D. Clinical and demographic predictors of Barrett's esophagus among patients with gastroesophageal reflux disease: a multivariable analysis in veterans. *J Clin Gastroenterol.* 2001;33:306–309. [PubMed: 11588545]
27. Ford AC, Forman D, Reynolds PD, Cooper BT, Moayyedi P. Ethnicity, gender, and socioeconomic status as risk factors for esophagitis and Barrett's esophagus. *Am J Epidemiol.* 2005;162:454–460. 10.1093/aje/kwi218. [PubMed: 16076833]
28. Corley DA, Kubo A, Zhao W. Abdominal obesity and the risk of esophageal and gastric cardia carcinomas. *Cancer Epidemiol Biomark Prev.* 2008;17:352–358. 10.1158/1055-9965.EPI-07-0748.
29. Abrams JA, Fields S, Lightdale CJ, Neugut AI. Racial and ethnic disparities in the prevalence of Barrett's esophagus among patients who undergo upper endoscopy. *Clin Gastroenterol Hepatol.* 2008;6:30–34. 10.1016/j.cgh.2007.10.006. [PubMed: 18063419]
30. Anderson LA, Watson RG, Murphy SJ, et al. Risk factors for Barrett's oesophagus and oesophageal adenocarcinoma: results from the FINBAR study. *World J Gastroenterol.* 2007;13:1585–1594. [PubMed: 17461453]
31. Corley DA, Kubo A, Levin TR, et al. Abdominal obesity and body mass index as risk factors for Barrett's esophagus. *Gastroenterology.* 2007;133:34–41. 10.1053/j.gastro.2007.04.046. [PubMed: 17631128]
32. Rubenstein JH, Mattek N, Eisen G. Age- and sex-specific yield of Barrett's esophagus by endoscopy indication. *Gastrointest Endosc.* 2010;71:21–27. 10.1016/j.gie.2009.06.035. [PubMed: 19748616]
33. Smith KJ, O'Brien SM, Green AC, Webb PM, Whiteman DC, Study of Digestive H. Current and past smoking significantly increase risk for Barrett's esophagus. *Clin Gastroenterol Hepatol.* 2009;7:840–848. 10.1016/j.cgh.2009.04.018. [PubMed: 19410015]

34. Spechler SJ, Jain SK, Tendler DA, Parker RA. Racial differences in the frequency of symptoms and complications of gastro-oesophageal reflux disease. *Aliment Pharmacol Ther.* 2002;16:1795–1800. [PubMed: 12269973]
35. Whiteman DC, Sadeghi S, Pandeya N, et al. Combined effects of obesity, acid reflux and smoking on the risk of adenocarcinomas of the oesophagus. *Gut.* 2008;57:173–180. 10.1136/gut.2007.131375. [PubMed: 17932103]
36. Le CT. A solution for the most basic optimization problem associated with an ROC curve. *Stat Methods Med Res.* 2006;15:571–584. 10.1177/0962280206070637. [PubMed: 17260924]
37. Rubenstein JH, Scheiman JM, Sadeghi S, Whiteman D, Inadomi JM. Esophageal adenocarcinoma incidence in individuals with gastroesophageal reflux: synthesis and estimates from population studies. *Am J Gastroenterol.* 2011;106:254–260. 10.1038/ajg.2010.470. [PubMed: 21139576]
38. Thrift AP, Kendall BJ, Pandeya N, Whiteman DC. A model to determine absolute risk for esophageal adenocarcinoma. *Clin Gastroenterol Hepatol.* 2013;11:138–144e2. 10.1016/j.cgh.2012.10.026. [PubMed: 23103823]
39. Xie SH, Lagergren J. A model for predicting individuals' absolute risk of esophageal adenocarcinoma: moving toward tailored screening and prevention. *Int J Cancer.* 2016;138:2813–2819. 10.1002/ijc.29988. [PubMed: 26756848]

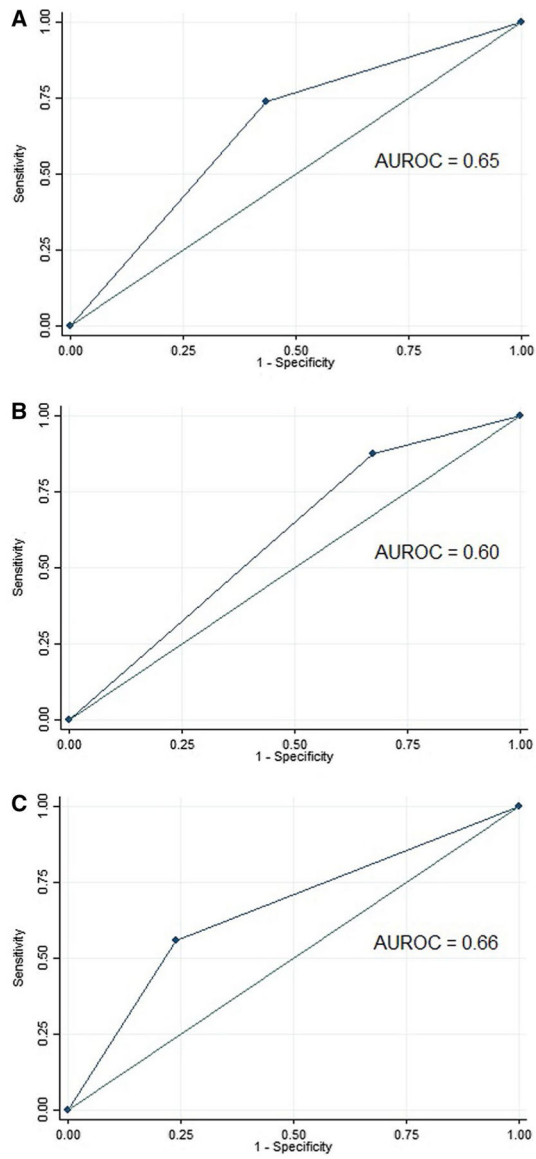


Fig. 1. Receiver operating characteristic curves for prespecified BE screening criteria among **a** all subjects, **b** men, and **c** women undergoing first-time upper endoscopy

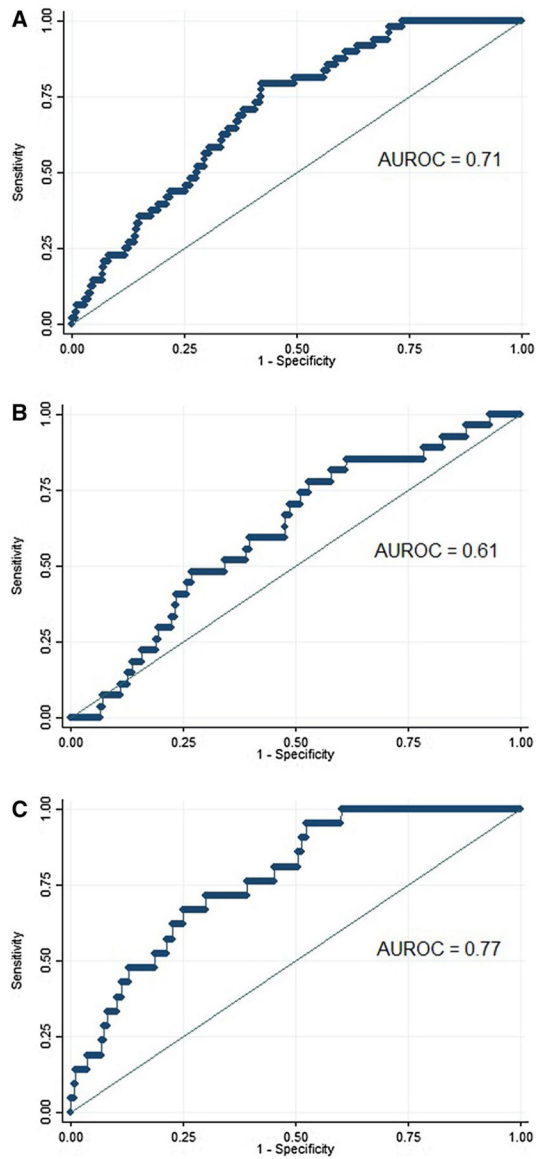


Fig. 2. Receiver operating characteristic curves for Barrett's esophagus risk score based on logistic regression modeling among **a** all subjects, **b** men, and **c** women undergoing first-time upper endoscopy. The model includes sex, gastroesophageal reflux disease, body mass index, and ever-smoker status, with sex removed from the models for sex-stratified analyses

Table 1
Patient characteristics and associations with the presence of BE on first upper endoscopy

	All subjects (n = 2931)	BE (n = 57)	Non-BE (n = 2874)	Odds ratio ^a (95% CI)	P
Age—mean (SD)	62.15 (7.2)	61.86 (6.7)	62.15 (7.2)	0.99 (0.96–1.03)	0.76
Male sex—no. (%)	1310 (44.7%)	32 (56.1%)	1278 (44.5%)	1.60 (0.94–2.71)	0.08
Race—no. (%)					0.26
Non-white	515 (17.6%)	6 (10.5%)	509 (17.7%)	1.00 (reference)	
White	1557 (53.1%)	36 (63.2%)	1521 (52.9%)	2.01 (0.84–4.79)	
Unknown	859 (29.3%)	15 (26.3%)	844 (29.4%)	1.50 (0.58–3.91)	
Ethnicity—no. (%)					0.94
Hispanic	511 (17.4%)	10 (17.5%)	501 (17.4%)	1.00 (reference)	
Non-hispanic	1442 (49.2%)	27 (47.4%)	1415 (49.2%)	0.96 (0.46–1.99)	
Unknown	978 (33.4%)	20 (35.1%)	958 (33.3%)	1.04 (0.49–2.25)	
BMI—mean (SD)	28.78 (7.5)	30.94 (9.6)	28.74 (7.5)	1.03 (1.00–1.07)	0.04
Smoking history—no. (%)					0.02
Never	1639 (55.9%)	23 (40.4%)	1616 (56.2%)	1.00 (reference)	
Ever	1067 (36.4%)	31 (54.4%)	1036 (36.1%)	2.10 (1.22–3.63)	
Unknown	225 (7.7%)	3 (5.3%)	222 (7.7%)	0.95 (0.28–3.19)	
GERD diagnosis—no. (%)	694 (23.7%)	24 (42.1%)	670 (23.3%)	2.39 (1.40–4.08)	0.002
GERD symptom—no. (%)	156 (5.3%)	3 (5.3%)	153 (5.3%)	0.99 (0.31–3.20)	1.00
PPI use—no. (%)	683 (23.3%)	22 (38.6%)	661 (23.0%)	2.10 (1.23–3.61)	0.01
H2RA use—no. (%)	197 (6.7%)	6 (10.5%)	191 (6.7%)	1.65 (0.70–3.90)	0.28
Family history—no. (%)	67 (2.3%)	2 (3.5%)	65 (2.3%)	1.57 (0.38–6.58)	0.38

BE Barrett’s esophagus, BMI body mass index, CI confidence interval, GERD gastroesophageal reflux disease, H2RA H2-receptor antagonist, PPI proton-pump inhibitor

^aUnadjusted odds ratios for association with BE

Performance characteristics of the prespecified BE screening criteria, in the full population and stratified by sex

Table 2

	Sensitivity (%)	Specificity (%)	Positive LR	Negative LR	AUROC
All Subjects ($n = 2931$)	73.68	56.68	1.70	0.46	0.65
Men ($n = 1310$)	87.50	32.55	1.30	0.38	0.60
Women ($n = 1621$)	56.00	76.00	2.33	0.58	0.66

AUROC area under the receiver operating characteristic curve, *BE* Barrett's esophagus, *LR* likelihood ratio

Table 3

Final BE prediction model for the full population based on EHR-derived patient data

	β coefficient	95% CI
GERD ^a	1.05	0.43 to 1.67
Male sex	0.52	– 0.07 to 1.12
Ever-smoker (vs. never)	0.66	0.07 to 1.25
BMI (per m/kg ²)	0.03	– 0.01 to 0.06

BE Barrett's esophagus, *BMI* body mass index, *CI* confidence interval, *EHR* electronic health record, *GERD* gastroesophageal reflux disease

^aGERD defined as GERD diagnosis or PPI use or H2RA use

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript