HHS Public Access

Author manuscript

Gastrointest Endosc. Author manuscript; available in PMC 2023 June 01.

Published in final edited form as:

Gastrointest Endosc. 2022 June ; 95(6): 1113–1122. doi:10.1016/j.gie.2021.12.034.

External validation of a model determining risk of neoplastic progression of Barrett's esophagus in a cohort of United States veterans

Theresa H. Nguyen, MD, MPH[1,2], **Aaron P. Thrift, PhD**[3,4], **Gyanprakash A. Ketwaroo, MD**[1], **Xianglin L. Du, PhD**[5], **Luis Leon Novelo, PhD**[6], **Rollin George, MD**[1], **Daniel G. Rosen, MD**[7,8], **Hashem B. El-Serag, MD, MPH**[1,2]

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

[2]Center for Innovations in Quality, Effectiveness and Safety (IQuESt), Michael E DeBakey Veterans Affairs Medical Center, Houston, Texas, USA

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

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

[5]Department of Epidemiology, The University of Texas School of Public Health, Houston, Texas, USA

[6]Department of Biostatistics and Data Science, The University of Texas Health Science Center at Houston - School of Public Health, Houston, Texas, USA

[7]Department of Pathology, Baylor College of Medicine, Houston, Texas, USA

[8]Department of Pathology, Michael E DeBakey Veterans Affairs Medical Center, Houston, Texas, USA

ADDRESS: Hashem B. El-Serag, MD, MPH, MEDVAMC 152, 2002 Holcombe Blvd., Houston, TX 77030, Phone (713) 798-647, hasheme@bcm.edu.

Author Involvement:

Theresa H. Nguyen: study concept and design, analysis and interpretation of data, statistical analysis, drafting of the manuscript, critical revision of the manuscript for important intellectual content

Aaron P. Thrift: study concept and design, analysis and interpretation of data, critical revision of the manuscript for important intellectual content

Xianglin L. Du: study concept and design, interpretation of data, critical revision of the manuscript for important intellectual content

Luis Leon Novelo: study concept and design, interpretation of data, critical revision of the manuscript for important intellectual content

Gyanprakash A. Ketwaroo: acquisition of data, critical revision of the manuscript for important intellectual content

Rollin George: acquisition of data, critical revision of the manuscript for important intellectual content

Daniel G. Rosen: acquisition of data, critical revision of the manuscript for important intellectual content

Hashem B. El-Serag: study concept and design, interpretation of data, critical revision of the manuscript for important intellectual content, obtained funding, study supervision

CONFLICT OF INTEREST: The authors report no personal or financial conflicts of interest.

Abstract

Background and Aims—Risk of esophageal adenocarcinoma (EA) in those with Barrett's esophagus (BE) is 11-fold greater than the general population. It remains unclear which BE patients are at highest risk of progression to EA. We aimed to validate a predictive model risk stratifying BE patients.

Methods—We conducted a retrospective cohort study at the Houston VA of consecutive patients with a new BE diagnosis from November 1990 to January 2019. Study follow-up was through 2/2020. Patients were excluded if they had no follow-up esophagogastroduodenoscopy (EGD) with esophageal biopsy after the initial BE-diagnosing EGD or evidence of high-grade dysplasia (HGD) or EA on initial EGD. We performed an external validation study of a risk model containing sex, smoking, BE length, and low-grade dysplasia (LGD) status and assessed discriminatory ability using area under the receiver operating characteristic curve (AUROC).

Results—Among 608 BE patients, 24 progressed to HGD/EA. The points-based model discriminated well with an AUROC of 0.72 (95% confidence interval [CI], 0.63–0.82). When categorized into low/intermediate/high-risk groups according to published cut-offs, the AUROC was poor at 0.57. Restructured into low-risk versus high-risk groups, the AUROC was 0.72 (95% CI, 0.64–0.80). Excluding baseline LGD did not reduce discriminatory ability (AUROC 0.73; 95% CI, 0.64–0.82).

Conclusion—This external validation provides further evidence that the model including sex, LGD status, smoking status, and BE length may help to risk stratify BE patients. A simplified version excluding LGD status and/or reducing the number of risk groups has increased utility in clinical practice without loss of discriminatory ability.

INTRODUCTION

Since the 1970s, the incidence of esophageal adenocarcinoma (EA) has been rising in developed countries by 3.5% to 8.0% per year.[1, 2] Despite increased screening and surveillance, EA patients continue to have a dismal average survival of 13 months.[3–6] Barrett's esophagus (BE), in which the normal esophageal squamous epithelium lining the lower end of the esophagus is replaced with specialized columnar intestinal epithelium, is the only known precursor to EA.[7] In the United States, 0.5% to 2% of the general adult population are estimated to have BE, and the risk of EA in those with BE is approximately 11-fold greater than that in the general population.[8] Clinical practice guidelines recommend surveillance every 3 to 5 years among all patients with nondysplastic BE (NDBE) to identify early neoplasia that can be cured with endoscopic resection.[9–11] However, given the low overall risk of neoplastic progression in NDBE (~0.5% per year), it remains unclear which patients with BE will develop EA and thus may benefit from endoscopic surveillance, resulting in oversurveillance, excess cost, and physical/emotional burden for many NDBE patients who are never likely to develop neoplasia.[12–15]

Reliable risk prediction models are needed to help tailor surveillance recommendations according to a patient's individual risk for neoplastic progression. Using data from a multicenter cohort of 2,697 BE patients (154 progressed to high grade dysplasia [HGD] or EA), Parasa et al[16] developed a mathematical model predicting the risk of neoplastic

progression in BE based on sex, smoking status, BE length, and presence of low-grade dysplasia (LGD) at the time of baseline endoscopy. This prediction model had good discriminatory ability in the internal validation cohort with an area under the receiver operating characteristic curve (AUROC) of 0.7. A clinically practical points-based system was subsequently developed based on the output of the same model, and used to stratify their cohort into 3 risk groups: low (0.13% annual risk of progression to HGD/EA), intermediate (0.73% annual risk), and high-risk (2.1% annual risk), without loss of discriminatory ability.

External validation of the Parasa et al scoring system[16] is needed in other BE populations because differences in population characteristics (such as distribution of risk factors) and risk of progression can affect the model's performance.[17] An external validation of this model was previously performed in the Northern Ireland Barrett's registry cohort (1,198 patients of whom 54 progressed to HGD/EA) with an AUROC close to that of the original study (0.68).[18] However, neither the original nor the validation study addressed the problematic concept of one-time confirmed LGD (2 pathologist reviews of a single biopsy) versus persistent LGD (LGD found on multiple endoscopies).[19, 20] In addition, the 3-risk category model can be difficult to interpret, as clinicians may have difficulty deciphering what to do with patients in the intermediate-risk group.[21] Therefore, the aim of this study was to externally validate the predictive model in a high risk Veteran population[22] and assess whether modifications to the model and its scoring system would improve the discriminatory performance or simplify the interpretation of the results.

METHODS

Study Design and Population

This was a retrospective cohort study of patients with BE at the Michael E. DeBakey Veteran Affairs Medical Center (MEDVAMC) in Houston, Texas, USA. We included consecutive patients with a new BE diagnosis on esophagogastroduodenoscopy (EGD) performed from November 1990 to January 2019 and confirmed on histopathological examination with specialized intestinal epithelium.[7] Patients were excluded from the analysis if they had (1) no follow-up EGD with esophageal biopsy after the initial BE-diagnosing EGD or (2) evidence of HGD or EA on the initial BE-diagnosing EGD. The study follow-up period was through the last date of medical record encounters or death through February 2020. During the study period, all endoscopists were expected to sample suspected and established BE with random biopsies every 1 to 2 cm in 4 quadrants with targeted biopsies of any areas suspicious for dysplasia. This research was approved by the Institutional Review Boards for Human Subjects Research for Baylor College of Medicine and the VA Research and Development Committee of the MEDVAMC (H-47857 approved August 18, 2020).

Data Collection and Measures

We performed a systematically structured manual review of the Veteran Affairs Computerized Patient Record System and abstracted the dates, endoscopic findings, and pathology results for every EGD and recorded these in the clinical database. Expert gastrointestinal pathologists examined all BE specimens during the study period. A second

dedicated expert gastrointestinal pathologist confirmed each diagnosis of dysplasia (LGD or HGD) or EA within BE as per routine clinical practice at the MEDVAMC.

The index EGD was defined as the first endoscopy with endoscopic and histologic BE. The primary outcome was neoplastic progression defined as HGD or EA on any follow-up endoscopy prior to any endoscopic treatment (ie, radiofrequency ablation, argon plasma coagulation, gold probe cautery, or endoscopic mucosal resection) following the initial BE-diagnosing EGD. Nonprogressors were those without neoplastic progression defined as no BE, NDBE, BE indefinite for dysplasia, or LGD BE on any follow-up EGD during the study period. Follow-up time started at the date of the index endoscopy and ended at the date of the first follow-up EGD demonstrating HGD/EA for progressors, last follow-up endoscopy before any endoscopic treatment for nonprogressors. Lost to followup was defined as those who did not receive a follow-up endoscopy within 5 years of NDBE or within 1 year of indefinite for dysplasia or LGD.[23]

We manually reviewed the medical records and extracted the following variables for the model by Parasa et al[16]: sex (male, female), smoking status (never, ever), BE length, and LGD status at index endoscopy. Using the Prague C and M criteria, we defined the length of BE as the highest value for circumference length (C) or the maximum (M) extent of the endoscopically visualized BE segment in centimeters (cm) at index endoscopy.[24]

We also extracted age at index endoscopy, race/ethnicity (non-Hispanic white, AfricanAmerican, and Hispanic), hiatal hernia at index endoscopy (absent or present), and height in inches and weight in pounds at time of index endoscopy. Body mass index (BMI) at index endoscopy was calculated by weight in pounds x 703/squared height in inches and stratified as normal (<25), overweight (≥25 and <30), and obese (≥30). Alcohol use was obtained from the most recent Alcohol Use Disorders Identification Test (AUDIT-C) screen prior to the index endoscopy (ever drinker: AUDIT-C > 4 in men and ≥ 3 in women). Proton pump inhibitor (PPI) use was obtained from VA pharmacy records and recorded as yes versus no.

Statistical Analysis

The mean follow-up time (and standard deviation [SD]) were calculated for the overall study cohort as well as separately for progressors and nonprogressors. Proportions of those lost to follow-up were calculated for nonprogressors who remained alive at the end of the study period. Using the points-based model developed by Parasa et al,[16] we assigned patients points based on length of BE at index endoscopy (1 point per 1 cm, up to maximum of 10 points), male sex (9 points), ever smoking (5 points), and presence of LGD on pathology at index endoscopy (11 points). Following the 3 risk-category based model by Parasa et al,[16] we stratified the cohort into 3 risk groups based on number of points assigned from: low-risk (0–10 points), intermediate-risk (11–20 points), and high-risk (>20 points) groups. We also simplified the risk-category based model by Parasa et al by dividing subjects into 2 risk groups based on points: low-risk (≤15 points) and high-risk (>15 points). This cutoff at 15 points was chosen, as 15 points was the halfway point of the intermediate group (11–20 points) of the 3 risk-category model.

We calculated the overall cumulative incidence and 95% confidence interval (CI) as well as annual risk of progression in the whole study cohort and stratified by the 3 risk groups based on the model by Parasa et al[16] and our simplified version with 2 risk groups. We performed competing risk Cox regression to compare cumulative incidence of progression per risk group.

We used Cox Proportional Hazards regression model to estimate hazard ratios (HR) and 95% CIs for the risk of neoplastic regression to HGD/EA and (1) the points-based model by Parasa et al,[16] (2) the 3 risk-category based model by Parasa et al.[16] (low-, intermediate-, and high-risk), (3) and our simplified 2 risk-category model (low-risk with ≤15 points and high-risk with >15 points). Model discrimination was quantified by calculating the AUROC and 95% CI of each model. Using a cutoff of low-risk for negative screening test and intermediate- and high-risk for positive screening test, we calculated the sensitivity, specificity, and corresponding 95% CIs. We also performed these models in a smaller cohort limited to only those diagnosed with BE from 2005 to 2019.

Comparison of Models with Additional Risk Factors and Existing Practice Guidelines—We performed bidirectional stepwise selection of candidate variables (eg, age, race/ethnicity, BMI, alcohol history, hiatal hernia on index endoscopy) to assess for potential additions to the model proposed by Parasa et al.[16] We also estimated the HRs and 95% CIs with Cox Proportional Hazards models of various existing practice guidelines: age, BE length, and LGD for the American College of Gastroenterology guidelines (ACG);[23] BE length ≥3 cm and LGD from the British Society of Gastroenterology (BSG);[25] and LGD only to mimic the American Gastroenterological Association (AGA) guidelines.[9] The AUROC and 95% CI for the discrimination of each model were assessed.

Sensitivity Analyses—We compared the AUROC for the model by Parasa et al.[16] to each of the following models: (1) excluding individual factors (sex, smoking status, BE length in cm, and low grade dysplasia status) one at a time from the model, (2) stratifying BE length by short (<3 cm) versus long (≥3 cm) rather than by each 1 cm, (3) restricting the definition of baseline LGD to only those with persistent LGD (LGD on 1 or more follow-up EGDs), (4) including LGD as an outcome of progression among those with NDBE or BE indefinite for dysplasia on index EGD, and (5) limiting the analysis to only those with index EGD performed prior to 2016 to allow longer follow-up time.

All analyses were performed using Stata/IC version 16.0 (StataCorp, College Station, Tex, USA), and a 2-tailed p-value of < 0.05 was considered statistically significant.

RESULTS

We identified 723 patients diagnosed with BE during 1990 to 2019, of whom 608 were included in the final analysis (Figure 1). The mean age of the study cohort was 61.6 years (standard deviation [SD], 8.6 years), and 95.9% (n=583) were male (Table 1). Most of the study cohort was non-Hispanic white (79.9%) followed by non-Hispanic black (10.2%) and Hispanic (9.9%). In the study cohort, 24 patients were progressors (development of HGD

or EA during follow-up), while 584 patients were nonprogressors (no BE, BE indefinite for dysplasia, or BE with LGD during follow-up).

The mean follow-up time after index endoscopy was 4.1 years (SD, 3.9 years) for the overall cohort, with longer average follow-up among nonprogressors 4.1 years (SD, 3.9 years) than progressors 3.5 years (SD, 3.4 years). There were no differences in losses to follow-up by the 3 risk groups in the model by Parasa et al[16]; 19.2% (n=15) were lost to follow up in the low-risk group, whereas 24.1% (n=82) of the intermediate-risk group and 36.2% (n=21) of the high-risk group were lost to follow-up (p=0.065).

The cumulative incidence of progression in the study cohort was 3.9% (95% CI, 2.7%5.8%), with an annual risk of 1.0% (95% CI, 0.6%−1.4%). Stratified by the proposed categories in the risk model by Parasa et al,[16] the cumulative incidence of progression was 1.1% (95% CI, 0.2%−7.8%) in the low-risk group, 4.2% (95% CI, 2.7%−6.5%) in the intermediate-risk group, and 6.1% (95% CI, 2.3%−15.2%) in the high-risk group (p-value by competing risk regression 0.326; Figure 1). The annual risk of progression was 0.34% in the low-risk group, 0.97% in the intermediate-risk group, and 1.5% in the high-risk group (p=0.351).

Stratified by the simplified output of 2 risk groups, the cumulative incidence of BE progression was 1.1% (95% CI, 0.4%−2.9%) in the low-risk group with ≤15 points, and 8.0% (95% CI, 5.2%−12.1%) in the high-risk group with >15 points (p-value by competing risk regression <0.001; Figure 3). The annual risk of progression was 0.3% in the low-risk group and 1.7% in the high-risk group.

Performance of the Model by Parasa et al[16]

In a Cox Proportional Hazards model fitting the 4 variables proposed by Parasa et al (sex, smoking status, BE length, and LGD status), only BE length was significantly associated with risk of progression (Table 2). The model with the 4 variables had an AUROC of 0.74 (95% CI, 0.64–0.84).

The point-based risk model proposed by Parasa et al was significantly associated with a 13% increase in risk of progression per additional point (HR 1.13; 95% CI, 1.05–1.23; Table 3), and a good discrimination with an AUROC of 0.72 (95% CI, 0.63–0.82). However, the 3-category model based on point cut-offs from Parasa et al had poor discriminatory ability with an AUROC of 0.57 (95% CI, 0.49–0.65), low specificity of 14.7% (95% CI, 12.0%−17.9%) at the intermediate-risk cut-off (>10 points), and low sensitivity of 16.7% (95% CI, 4.7%−37.4%) at the high-risk cut-off (>20 points). Conversely, our proposed re-categorization with 2 groups (low-risk 0–15 points, high-risk >15 points) was associated with a 5.89 times increased risk of progression (95% CI, 2.01–17.3) in the highrisk group compared with the low-risk group, with good discrimination (AUROC 0.72, 95% CI, 0.64–0.80), and high specificity (60.8%) and sensitivity (83.3%). The findings were similar when limiting the cohort for those diagnosed with BE from 2005 to 2019 (n=545) with good discrimination by the points-based risk model and recategorized 2-risk group model and poor discrimination by the 3-category model.

Models with Additional Variables and Existing Practice Guidelines

Age ≥70 years (ref <60 years: adjusted HR 7.35, 95% CI, 2.00–27.2) was associated with risk of progression to HGD or EA in a larger model that included sex, smoking status, BE length, LGD status, age, and BMI variables (Table 2). However, there was no improvement in discrimination by the (AUROC 0.73 0.62–0.84) compared to the model by Parasa et al.[16] s (predictor variables: age, BE length, and LGD at index endoscopy) and the BSG guidelines[25] (predictor variables: BE length and LGD at index endoscopy) had similar discriminatory ability compared with the model by Parasa et al[16] with an AUROC of 0.72 (Table 2). However, the model proposed by AGA guidelines[9] containing only LGD status at index endoscopy as a predictor variable had poor discriminatory ability with an AUROC of 0.52 (95% CI, 0.46–0.58).

Sensitivity Analyses

In the first set of sensitivity analyses, we excluded individual predictive variables one at a time from the points-based model by Parasa et al, and found that each permutated model significantly predicted risk of progression with good discrimination except for the model excluding BE length (AUROC 0.51, 95% CI, 0.41–0.61; Table 4). Notably, the points-based model excluding LGD status at index endoscopy performed similar to the full points-based model with an AUROC of 0.73 (95% CI, 0.64–0.82).

In the second set of sensitivity analyses, the model containing variables proposed by Parasa et al with BE length variable recategorized as <3 cm and ≥3 cm performed well with an AUROC of 0.73 (95% CI, 0.63–0.83). This was similar to the original model with continuous BE length (Table 2).

Of the 25 BE patients with confirmed LGD on index endoscopy, 11 had persistent LGD (defined as LGD on 1 or more follow-up EGDs). In the third set of sensitivity analyses, we examined the models using only baseline LGD that remained persistent LGD. There was no statistically significant improvement in the discrimination of both the point-based risk score model (AUROC 0.74; 95% CI, 0.64–0.83) or the simplified 2-risk category model (AUROC 0.73; 95% CI, 0.65–0.81) compared with the models containing confirmed LGD (AUROC 0.72).

In the fourth sensitivity analysis that included LGD as an outcome of progression, 93 developed LGD, HGD, or EAC from a cohort of 583 with NDBE or BE indefinite for dysplasia on index EGD, with an annual risk of progression of 3.8% (95% CI, 3.1% −4.7%). There was no statistically significant improvement in discrimination between the point-based risk score model predicting progression to LGD, HGD, or EAC (AUROC 0.71, 95% CI, 0.66–0.77) and the point-based model containing LGD as a predictor of progression to HGD or EAC (AUROC 0.72).

When limiting the analysis to only those with index EGD performed before 2016, 23 (4.2%) of 552 patients progressed during the follow-up period. There was slight improvement in the AUROC to 0.75 (95% CI, 0.67–0.83) compared with the model containing patients from the entire study period through 2019 (AUROC 0.72).

DISCUSSION

In a cohort of 608 patients diagnosed with BE from 1990 to 2019 at the MEDVAMC, we externally validated the risk prediction model from Parasa et al[16], and supported the findings that clinical factors based on sex, smoking status, length of BE at diagnosis, and presence of LGD at time of BE diagnosis could help to risk stratify patients with BE for progression to advanced neoplasia. The risk model output had an AUROC of 0.72 (95% CI, 0.63–0.82) in our study population, which was similar to the internal validation results reported by Parasa et al (AUROC 0.70).[16]

However, the 3-risk group categorization of the model output proposed by Parasa et al.[16] failed to correctly risk stratify patients for progression in our study cohort with an AUROC of 0.57. When we restructured the output of the points-based model into 2 groups (lowrisk with ≤15 points and high-risk with >15 points), the cumulative incidence was 1.1% in the low-risk group and 8.0% in the high-risk group, and the model AUROC was good at 0.72 (95% CI, 0.64–0.80). Furthermore, clinicians may have difficulty making recommendations for patients who fall in the intermediate-risk group of a 3-risk group model. Hence, a simplified risk model format with 2 groups (low-risk and high-risk) was able to predict risk of neoplastic progression in our study population and is more appropriate for clinical decision-making.[21]

The discriminatory ability of the risk score model without LGD status at baseline was not different than the original full risk prediction model containing LGD status (AUROC 0.73 vs 0.72, respectively). Similarly, the model mimicking the AGA guidelines, which uses LGD status only, had a poor prediction of progression with an AUROC of 0.52. This finding may be due to variability in diagnostic accuracy of LGD by pathologists; a multi-center study from 2017 showed poor interobserver agreement for LGD even among expert GI pathologists.[19] *Persistent* LGD status may be a better predictor of neoplastic progression, in which endoscopic ablative therapy is preferred over surveillance.[23] In our study, the models requiring persistent LGD had improved discriminative ability compared to those with one-time confirmed LGD, although this difference did not reach statistical significance. Thus, a simpler model without confirmed LGD status may perform better in predicting progression using data from only the index endoscopy, however the inclusion of persistent LGD status over several endoscopies may improve neoplastic risk prediction overall. Further, one can argue that LGD should be classified as an outcome event, given that endoscopic treatment with complete eradication of BE is recommended. In this study, we found that the model containing gender, BE length, and smoking status was able to adequately predict progression to LGD, HGD, or EAC (AUROC 0.71).

Our models in which we simplified BE length variable to short (<3 cm) and long (≥3 cm) had no change in predictive ability compared with BE length per cm (AUROC 0.73 vs 0.74, respectively). A previous validation study of the Prague C&M classification in clinical practice found that absolute agreement of BE length within ≤1 cm was only 68%, with no improvement in BE experts compared with community hospital endoscopists.[26] Classifying BE length by short and long would reduce the possibility for misclassification without compromising the predictive ability of the model to predict risk of progression.

Few prediction models for the future risk of EA developing in BE have been developed, and even less have been validated in external populations.[16, 27] External validation in multiple populations is an important step in the development of prediction models, as it addresses transportability to different populations temporally and geographically.[21] Ours is the first study to perform an external validation of the model by Parasa et al in a United States population, as well as in a Veteran-only population which has a higher prevalence of BE and EA than the general population due to the disproportionate number of older, male smokers.[22, 28] Other strengths of our study included the use of a consecutively and prospectively maintained cohort of BE patients, and our strict BE diagnosis definition. Additionally, our study demonstrated already known associations between risk of progression and BE length, conferring internal validity to our findings.

Our study had several limitations. It was conducted in a Veteran population of mainly non-Hispanic white men, limiting generalizability of the results to other populations. There was a potential for misclassification of the predictor variables (smoking status, LGD status, and BE length); however, the performance of the model was robust despite exclusion of each risk factor except BE length. There may have been misclassification of the outcome (ie, progression to HGD or EA) due to loss of follow-up, as the mean follow-up period was only 4.1 years; however, those in the low-risk group had shorter follow-up duration (3.7 years) than the overall cohort. Further, adherence to biopsy protocol was not reliably recorded and could be a cause for misclassification of dysplasia status. Despite strict inclusion definition, approximately 22% had no BE on follow-up EGD. We suspect that these cases had short or very short BE length, especially those with hiatal hernias where sampling is known to be difficult and inconsistent. Last, the sample size may have been too small to detect differences between progressors and nonprogressors, and so the addition of new variables, such as age, to the model was not possible.

In summary, we conducted the first external validation and modification of the risk prediction model for neoplastic progression in BE by Parasa et al in a United States Veteran population. The model including sex, LGD status, smoking status, and BE segment length performed well in our study population. Further, we found similar performance with our simplified versions of the model with (1) BE length categorized as <3 cm and ≥3 cm, (2) excluding LGD status from the model, and (3) reducing the number of risk groups to low- and high-risk; the simplified models may have increased utility in clinical practice. Future studies are required including prospective evaluation of the model in a cohort with strict inclusion/exclusion criteria, as well as improvement in the precision of the model with the identification and addition of genetic risk factors for neoplastic progression.

Acknowledgments

GRANT SUPPORT: This work is funded in part by National Institutes of Health grant NCI R01 116845, and the Texas Digestive Disease Center NIH DK58338. Dr. El-Serag is also supported by NIDDK K24–04-107. Dr. Nguyen is supported by NIH T-32 5T32DK083266–07 Grant. This research was supported in part with resources at the VA HSR&D Center for Innovations in Quality, Effectiveness and Safety (#CIN 13–413), at the Michael E. DeBakey VA Medical Center, Houston, TX. The opinions expressed reflect those of the authors and not necessarily those of the Department of Veterans Affairs, the US government or Baylor College of Medicine.

REFERENCES

1. Cook MB, Thrift AP. Epidemiology of Barrett's esophagus and esophageal adenocarcinoma: Implications for screening and surveillance. Gastrointest Endosc Clin N Am 2021;31:1–26. [PubMed: 33213789]
2. Thrift AP. Global burden and epidemiology of Barrett oesophagus and oesophageal cancer. Nat Rev Gastroenterol Hepatol 2021.
3. Edgren G, Adami HO, Widerpass E, et al. A global assessment of the oesophageal adenocarcinoma epidemic. Gut 2013;62:1406–14. [PubMed: 22917659]
4. Thrift AP, Whiteman DC. The incidence of esophageal adenocarcinoma continues to rise: analysis of period and birth cohort effects on recent trends. Ann Oncol 2012;12:3155–62.
5. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worlwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424. [PubMed: 30207593]
6. Haiyu Z, Xiaofeng P, Xiangquiong M, et al. Incidence and survival changes in patients with esophageal adenocarcinoma during 1984–2013. Biomed Res Int 2019:7431850.
7. Spechler SJ. Barrett's esophagus and esophageal adenocarcinoma: pathogenesis, diagnosis, and therapy. Med Clin North Am 2002;86:1423–45. [PubMed: 12510459]
8. Shaheen NJ, Richter JE. Barrett's oesophagus. Lancet 2009;373:850–61. [PubMed: 19269522]
9. Spechler SJ, Sharma P, Souza RF, et al. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. Gastroenterology 2011;140:1084–91. [PubMed: 21376940]
10. Sharma P, Shahee NJ, Katzka D, et al. . AGA clinical practice update on endoscopic treatment of Barrett's esophagus with dysplasia and/or early cancer: Expert review. Gastroenterology 2020;158:760–9. [PubMed: 31730766]
11. El-Serag HB, Naik AD, Duan Z, et al. Surveillance endoscopy is associated with improved outcomes of oesophageal adenocarcinoma detected in patients with Barrett's oesophagus. Gut 2016;65.
12. Hvid-Jensen F, Pedersen L, Drewes AB, et al. Incidence of adenocarcinoma among patients with Barrett's esophagus. N Engl J Med 2011;365:1375–83. [PubMed: 21995385]
13. Crockett SD, Lipkus IM, Bright SD, et al. Overutilization of endoscopic surveillance in nondysplastic Barrett's esophagus: a multicenter study. Gastrointest Endosc 2012;75:23–31. [PubMed: 22100301]
14. Saxena N, Inadomi JM. Effectiveness of cost-effectiveness of endoscopic screening and surveillance. Gastroitnest Endosc Clin N Am 2017;27:397–421.
15. Kruijshaar ME, Kerkshof M, Siersema PD, et al. The burden of upper gastrointestinal endoscopy in patients with Barrett's esophagus. Endoscopy 2006;38:873–8. [PubMed: 17019759]
16. Parasa S, Vennalganti S, Gaddam S, et al. Development and validation of a model to determine risk of progression of Barrett's esophagus to neoplasia. Gastroenterology 2018;154:1282–9. [PubMed: 29273452]
17. Thrift AP, Kanwal F, El-Serag HB. Prediction models for gastrointestinal and liver diseases: Too many developed, too few validated. Clin Gastroenterol Hepatol 2016;14:1678–80. [PubMed: 27574756]
18. Kunzmann AT, Thrift AP, Johnston BT, et al. External validation of a model to determine risk of progression of Barrett's oesophagus to neoplasia. Aliment Pharmacol Ther 2019;49:1274–81. [PubMed: 30950101]
19. Vennalaganti P, Kanakadandi V, Golblum JR, et al. Discordance among pathologists in the United States and Europe in diagnosis of low-grade dysplasia for patients with Barrett's esophagus. Gastroenterology 2017;152:564–70. [PubMed: 27818167]
20. Thota PN, Lee HJ, Goldblum JR, et al. Risk stratification of patients with Barrett's esophagus and low-grade dysplasia or indefinite for dysplasia. Clin Gastroenterol Hepatol 2015;13:459–65. [PubMed: 25102445]

21. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: Seven steps for development and an ABCD for validation. Eur Heart J 2014;35:1925–31. [PubMed: 24898551]
22. Nguyen TH, Thrift AP, Rugge M, et al. Prevalence of Barrett's esophagus and performance of societal screening guidelines in an unreferred primary care population of U.S. veterans. Gastrointest Endosc 2020.
23. Shaheen NJ, Falk GW, Iyer PG, et al. ACG clinical guidelines: Diagnosis and management of Barrett's esophagus. Am J Gastroenterol 2016;111:30–50. [PubMed: 26526079]
24. Sharma P, Dent J, Armstrong D, et al. The development and validation of an endoscopic grading system for Barrett's esophagus: the Prague C & M criteria. Gastroenterology 2006;131:1392–9. [PubMed: 17101315]
25. Fitzgeral RC, di Pietr M, Ragunath K, et al. . British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. Gut 2014;63:7–42. [PubMed: 24165758]
26. Alvarez Herrero L, Curvers WL, van Vilsteren FG, et al. Validation of the Prague C&M classification of Barrett's esophagus in clinical practice. Endoscopy 2013;45:876–82. [PubMed: 24165812]
27. Sato F, Jin Z, Schulmann K, et al. Three-tiered risk stratification model to predict progression in Barrett's esophagus using epigenetic and clinical features. PLoS One 2008;3:e1890.
28. Connor MJ, Weston AP, Mayo MS, et al. The prevalence of Barrett's esophagus and erosive esophagitis in patients undergoing upper endoscopy for dyspepsia in a VA population. Dig Dis Sci 2004;49:920–4. [PubMed: 15309878]

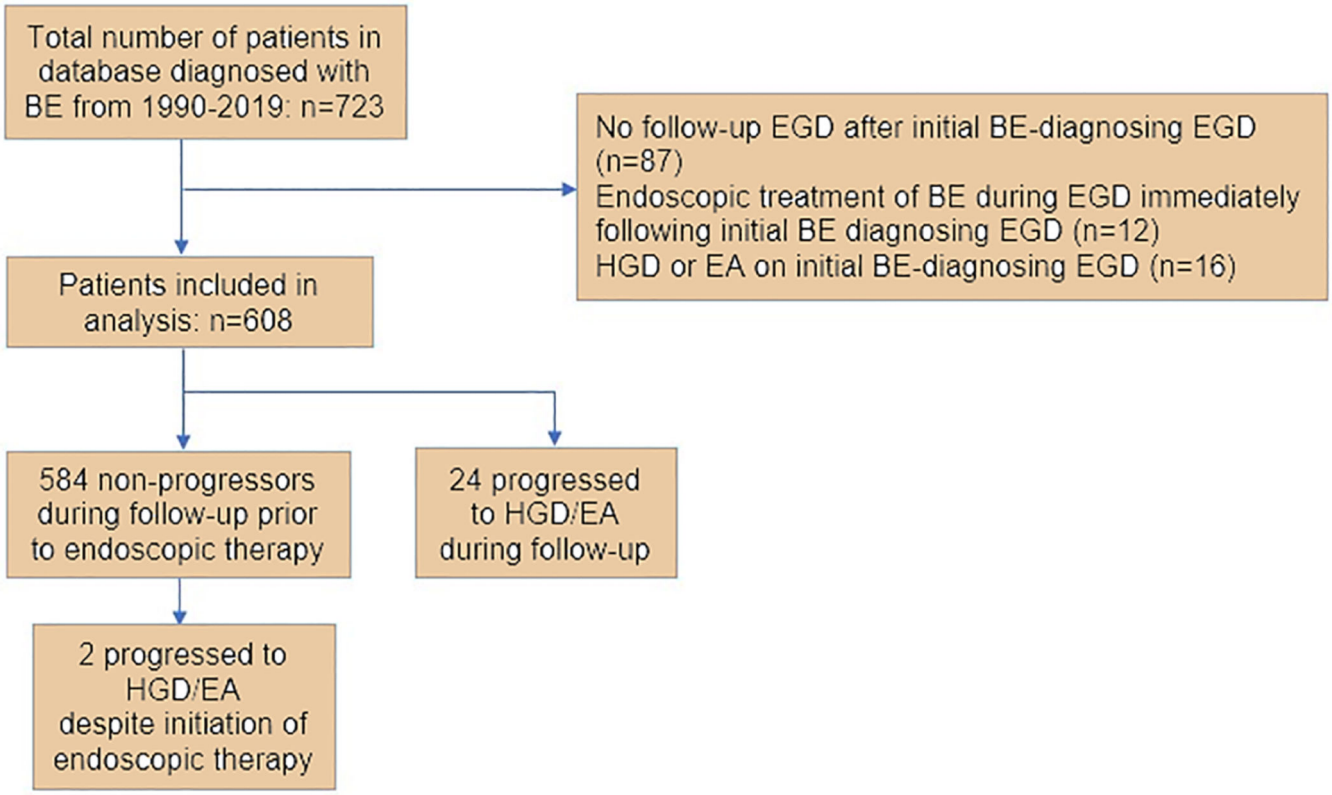

Figure 1.

Flow diagram of the 723 patients diagnosed with Barrett's esophagus (BE) at the Michael E. DeBakey Veteran Affairs Medical Center from 1990 to 2019.

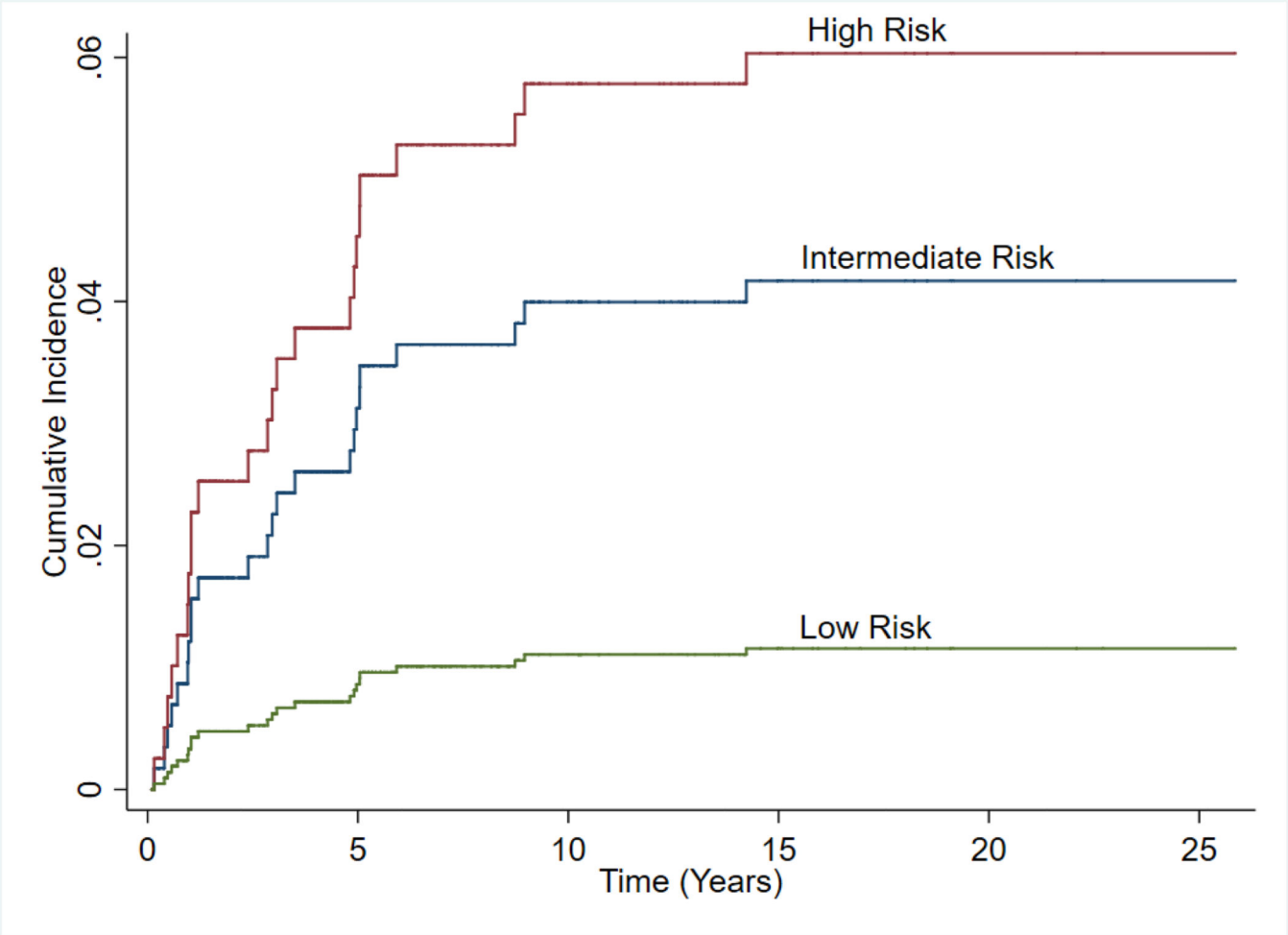

Figure 2.

Cumulative incidence of progression to high-grade dysplasia or esophageal adenocarcinoma by the 3 risk groups proposed by Parasa et al.[21] (green=low risk, blue=intermediate risk, red=high risk) were not different (p-value by competing-risks regression 0.326).

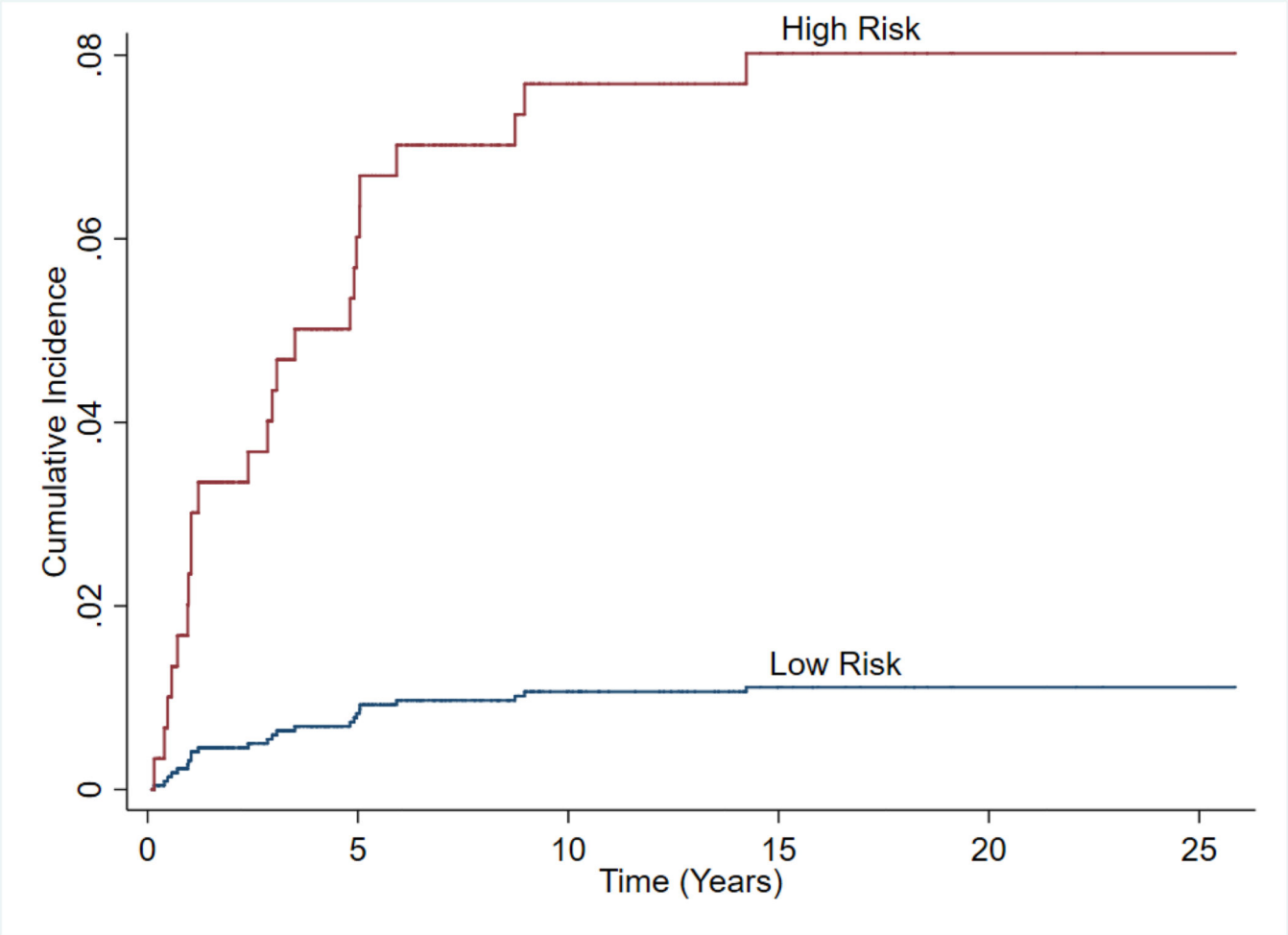

Figure 3.

Cumulative incidence of progression to high grade dysplasia or esophageal adenocarcinoma by the points-based risk model proposed by Parasa et al.[16] re-categorized into low risk with ≤15 points (blue) and high risk with >15 points (red) in a competing-risks regression (p-value<0.001).

Table 1.

Sociodemographic and clinical characteristics of the entire study cohort (n=608) and stratified by nonprogressors (n=584) and progressors (n=24) defined as development of high-grade dysplasia or esophageal adenocarcinoma during follow-up.

	Study cohort n (%)	Nonprogressors n (%)	Progressors n (%)
Frequency	**608**	**584**	**24**
Age in years			
<60	211 (34.7)	206 (35.3)	5 (20.8)
60 to <70	323 (53.1)	309 (52.9)	14 (58.4)
70+	74 (12.2)	69 (11.8)	5 (20.8)
Race			
NH White	486 (79.9)	466 (79.8)	20 (83.3)
NH Black	62 (10.2)	61 (10.5)	1 (4.2)
Hispanic	60 (9.9)	57 (9.8)	3 (12.5)
Sex			
Male	583 (95.9)	559 (95.7)	24 (100.0)
Female	25 (4.1)	25 (4.3)	0 (0.0)
BMI			
<25	109 (17.9)	106 (18.1)	3 (12.5)
25 to <30	234 (38.5)	227 (38.9)	7 (29.2)
30+	265 (43.6)	251 (43.0)	14 (58.3)
Smoking Status			
Never smoker	156 (25.7)	148 (25.3)	8 (33.3)
Ever smoker	452 (74.3)	436 (74.7)	16 (66.7)
Alcohol Drinking Status			
Never Drinker	253 (41.6)	242 (41.4)	11 (45.8)
Ever Drinker	355 (58.4)	342 (58.6)	13 (54.2)
PPI use			
No	186 (30.6)	180 (30.8)	6 (25.0)
Yes	422 (69.4)	404 (69.2)	18 (75.0)
Hiatal Hernia			
Absent	185 (30.4)	178 (30.5)	7 (29.2)
Present	423 (69.6)	406 (69.5)	17 (70.8)
BE Length in cm (mean, SD)	2.8 (3.0)	2.7 (3.0)	5.2 (3.5)
BE Length			
<3 cm	401 (65.9)	395 (67.6)	6 (25.0)
>3 cm	207 (34.1)	189 (32.4)	18 (75.0)
Dysplasia Status at Baseline Endoscopy			
None	457 (75.2)	440 (75.3)	17 (70.8)
Indefinite	126 (20.7)	121 (20.7)	5 (20.8)
LGD	25 (4.1)	23 (3.9)	2 (8.4)
Persistence of LGD at Baseline Endoscopy			

	Study cohort n (%)	Nonprogressors n (%)	Progressors n (%)
Frequency	**608**	**584**	**24**
No	597 (98.2)	575 (98.5)	22 (91.7)
Yes	11 (1.8)	9 (1.5)	2 (8.3)
Risk Score from Parasa Model[16] (mean, SD)	15.7 (4.7)	15.6 (4.6)	18.5 (4.7)
3 Risk Categories from Parasa Model [16]			
Low Risk (0–10 points)	87 (14.3)	86 (14.7)	1 (4.2)
Intermediate Risk (11–20 points)	455 (74.8)	436 (74.7)	19 (79.2)
High Risk (21 + points)	66 (10.9)	62 (10.6)	4 (16.7)
Simplified 2 Risk Categories			
Low Risk (<15 points)	59 (59.1)	355 (60.8)	4 (16.7)
High Risk (>15 points)	249 (40.9)	229 (39.2)	20 (83.3)

Abbreviations: SD=Standard deviation; NH=Non-Hispanic; BMI=body mass index; BE=Barrett's esophagus; cm=centimeters; LGD=low-grade dysplasia

Table 2.

Hazard ratio (HR) estimating risk of progression to high grade dysplasia or esophageal adenocarcinoma and area under the receiver operating characteristic curve (AUROC) in a Cox Proportional Hazards regression model containing predictive variables proposed by Parasa et al.

	Adjusted HR (95% CI)	AUROC (95% CI)
Variables in Model by Parasa et al.		0.74 (0.64–0.84)
Male sex (ref: female) *	--	
Ever smoker (ref: never smoker)	0.89 (0.37–2.14)	
BE length (per 1cm)	1.17 (1.06–1.29)	
Low grade dysplasia (ref: no LGD)	3.18 (0.74–13.7)	
Model from Bidirectional Stepwise Selection		0.73 (0.62–0.84)
Male sex (ref: female) *	--	
Ever smoker (ref: never smoker)	0.99 (0.41–2.40)	
BE length (per 1cm)	1.15 (1.05–1.26)	
LGD status (ref: no LGD)	2.45 (0.56–10.78)	
Age (ref: <60 years)	ref	
60 to <70 years	2.60 (0.90–7.49)	
70+ years	7.35 (2.00–27.2)	
BMI (ref: <25)	ref	
25 to <30	1.08 (0.28–4.24)	
30+	2.47 (0.68–9.06)	
Variables from ACG Guidelines		0.72 (0.61–0.83)
Age in years	1.07 (1.02–1.13)	
BE length in cm	1.17 (1.07–1.29)	
Low grade dysplasia (ref: no LGD)	2.96 (0.69–12.7)	
Variables from BSG Guidelines		0.72 (0.62–0.81)
BE length >3cm (Ref <3cm)	4.48 (1.76–11.4)	
Low grade dysplasia (ref: no LGD)	2.55 (0.59–10.9)	
Variables from AGA Guidelines		0.52 (0.46–0.58)
Low grade dysplasia (ref: no LGD)	3.29 (0.77–14.1)	

Abbreviations: HR=Hazard ratio; CI=confidence interval; AUROC=area under receiver operating characteristic; ref=reference; BE=Barrett's esophagus; LGD=low-grade dysplasia; NH=non- Hispanic; BMI=body mass index

* numbers of females too low for comparison

Table 3.

Hazard ratio (HR) estimating risk of progression to high grade dysplasia or esophageal adenocarcinoma, area under the receiver operating characteristic curves (AUROC), sensitivity, specificity, and corresponding 95% confidence intervals (CI) are reported for variations of the points-based model predicting risk of progression proposed by Parasa et al.

	Hazard ratio (95% CI)	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Full Cohort				
Parasa risk score (per one point)	1.13 (1.05–1.23)	0.72 (0.63–0.82)	--	--
Parasa Risk Categories (ref low risk: 0–10 points)		0.57 (0.49–0.65)	--	--
Intermediate risk: 11–20 points	2.93 (0.39–22.0)	--	95.8 (78.9–99.9)	14.7 (12.0–17.9)
High risk: 21+ points	4.75 (0.53–42.6)	--	16.7 (4.7–37.4)	89.4 (86.6–91.8)
Parasa risk score >15 points (ref: <15 points)	5.89 (2.01–17.3)	0.72 (0.64–0.80)	83.3 (62.6–95.3)	60.8 (56.7–64.8)
Cohort from 2005–2019				
Parasa risk score (per one point)	1.13 (1.05–1.23)	0.74 (0.65–0.84)		
Parasa Risk Categories (ref low risk: 0–10 points)		0.59 (0.49–0.68)		
Intermediate risk: 11–20 points	2.85 (0.37–21.56)		95.0 (75.1–99.9)	15.6 (12.6–19.0)
High Risk: 21+ points	4.57 (0.52–41.9)		20.0 (5.7–43.7)	89.3 (86.4–91.8)
Parasa risk score >15 points (ref: <15 points)	8.02 (2.34–27.4)	0.74 (0.66–0.83)	85.0 (62.1–96.8)	63.4 (59.1–67.6)

Abbreviations: BE = Barrett's esophagus; LGD=low-grade dysplasia; CI=confidence interval; AUROC=area under the receiver operating characteristic

Table 4.

The hazard ratios, 95% confidence intervals, and related area under the receiver operating characteristic curve estimating risk of progression to HGD/EA are reported for the following sensitivity analyses: 1) excluding individual factors one at a time from the point-based model by Parasa et al.,[16] 2) stratifying BE length by <3cm versus ≥ 3cm in the model by Parasa et al. rather than by each 1 cm, 3) with persistent LGD status (LGD on 1 or more follow-up EGDs) rather than confirmed LGD status (agreement by 2 pathologists of LGD on single EGD), and 4) classifying LGD as an outcome of progression with HGD and EAC.

	Hazard Ratio (95% CI)	AUROC (95% CI)
Permutations of Parasa points-based model		
Excluding Sex	1.13 (1.04–1.23)	0.71 (0.62–0.81)
Excluding Smoking	1.15 (1.07–1.24)	0.74 (0.66–0.83)
Excluding BE Length	1.05 (0.92–1.19)	0.51 (0.41–0.61)
Excluding LGD	1.14 (1.03–1.26)	0.73 (0.64–0.82)
Parasa model with categorized BE length		0.73 (0.63–0.83)
Male sex (ref: female)*	--	
Ever smoker (ref: never)	0.78 (0.33–1.83)	
Low grade dysplasia (ref: no LGD)	2.52 (0.59–10.8)	
BE length >3cm (ref: <3cm)	4.32 (1.69–11.0)	
Reclassifying LGD as p*ersistent* LGD		
Parasa Risk Score (per one point)	1.16 (1.07–1.25)	0.74 (0.64–0.83)
Parasa Risk Categories (ref Low Risk: 0–10 points)		0.58 (0.50–0.66)
Intermediate Risk: 11–20 points	2.94 (0.39–22.1)	
High Risk: 21+ points	5.51 (0.61–49.4)	
Parasa Risk Score >15 points (ref: <15 points)	6.13 (2.09–18.0)	0.73 (0.65–0.81)
Classifying LGD as outcome		
Parasa Risk Score (excluding LGD as predictor)	1.14 (1.09–1.20)	0.71 (0.66–0.77)

Abbreviations: AUROC=area under receiver operating characteristic curve; LGD=low- grade dysplasia; BE=Barrett's esophagus; cm=centimeters; HGD=high-grade dysplasia; EA=esophageal adenocarcinoma

*numbers of females too low for comparison