Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2017 February ; 26(2): 240–248. doi: 10.1158/1055-9965.EPI-16-0640.

A Tissue Systems Pathology Test Detects Abnormalities Associated with Prevalent High Grade Dysplasia and Esophageal Cancer in Barrett's Esophagus

Rebecca J. Critchley-Thorne[1,*], **Jon M. Davison**[2], **Jeffrey W. Prichard**[3], **Lia M. Reese**[1], **Yi Zhang**[1], **Kathleen Repa**[1], **Jinhong Li**[3], **David L. Diehl**[3], **Nirag C. Jhala**[4,‡], **Gregory G. Ginsberg**[4], **Maureen DeMarshall**[4], **Tyler Foxwell**[2], **Blair A. Jobe**[5], **Ali H. Zaidi**[5], **Lucas C. Duits**[6], **Jacques J.G.H.M. Bergman**[6], **Anil Rustgi**[4], and **Gary W. Falk**[4]

[1]Cernostics, Inc., 235 William Pitt Way, Pittsburgh, PA [2]Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA [3]Department of Laboratory Medicine, Geisinger Medical Center, Danville, PA [4]Division of Gastroenterology, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA [5]Esophageal and Lung Institute, Allegheny Health Network, Pittsburgh, Pennsylvania, PA [6]Department of Gastroenterology and Hepatology, Academic Medical Centre, Amsterdam, Netherlands

Abstract

Background—There is a need for improved tools to detect high grade dysplasia (HGD) and esophageal adenocarcinoma (EAC) in patients with Barrett's esophagus (BE). In previous work, we demonstrated that a 3-tier classifier predicted risk of incident progression in BE. Our aim was to determine if this risk classifier could detect a field effect in non-dysplastic (ND), indefinite for dysplasia (IND) or low-grade dysplasia (LGD) biopsies from BE patients with prevalent HGD/EAC.

Methods—We performed a multi-institutional case-control study to evaluate a previously developed risk classifier that is based upon quantitative image features derived from 9 biomarkers and morphology, and predicts risk for HGD/EAC in BE patients. The risk classifier was evaluated in ND, IND and LGD biopsies from BE patients diagnosed with HGD/EAC on repeat endoscopy (prevalent cases, n=30, median time to HGD/EAC diagnosis 140.5 days) and non-progressors (controls, n=145, median HGD/EAC-free surveillance time 2,015 days).

Results—The risk classifier stratified prevalent cases and non-progressor patients into low-, intermediate- and high-risk classes (odds ratio, 46.0; 95% confidence interval, 14.86–169 (high-

[*]**Correspondence:** Rebecca Critchley-Thorne, Ph.D., Cernostics, Inc., 235 William Pitt Way, Pittsburgh, PA 15238., rthorne@cernostics.com, Tel/Fax: 412-828-0900.

[‡]Currently at: Department of Pathology and Laboratory Medicine, Temple University School of Medicine, Philadelphia, PA.

Potential competing interests: Lia Reese, Anil Rustgi and Rebecca Critchley-Thorne hold equity ownership or stock options in Cernostics, Inc., the commercial entity that developed the proprietary TissueCypher technology used in this study. Lia Reese, Kathleen Repa, Rebecca Critchley-Thorne are current or former full-time employees of Cernostics, Inc. Yi Zhang is a paid consultant to Cernostics, Inc. Rebecca Critchley-Thorne reports being an inventor on patents and patent applications for the TissueCypher® technology used in this study. Jeffrey Prichard, Jinhong Li and David Diehl are employees of Geisinger Health System, which holds a financial interest in Cernostics, Inc.

risk vs low-risk); p<0.0001). The classifier also provided independent prognostic information that outperformed the subspecialist and generalist diagnosis.

Conclusion—A tissue systems pathology test better predicts prevalent HGD/EAC in BE patients than pathologic variables. The results indicate that molecular and cellular changes associated with malignant transformation in BE may be detectable as a field effect using the test.

Impact—A tissue systems pathology test may provide an objective method to facilitate earlier identification of BE patients requiring therapeutic intervention.

Keywords

Barrett's esophagus; prevalent; biomarkers; digital pathology; esophageal adenocarcinoma

Introduction

Barrett's esophagus (BE) is a precursor to esophageal adenocarcinoma (EAC), which is the fastest growing cancer type by incidence in the US with 5 year survival rates of 18% (1). EAC can be prevented if dysplasia is detected and treated early with endoscopic therapies such as radiofrequency ablation (RFA) and/or endoscopic mucosal resection (EMR) (2–4). Current guidelines from the American College of Gastroenterology (ACG) recommend surveillance by endoscopy with biopsies at intervals determined by the pathologic diagnosis (5). The diagnosis of dysplasia in BE is limited by the random nature of endoscopic sampling, which may miss dysplastic areas, and by inter-observer variation (6). While subtle lesions containing high grade dysplasia (HGD) and EAC can be detected by expert endoscopists at high-volume centers, recognition of subtle lesions is more challenging in the community setting (7). These limitations can result in repeat endoscopies and delayed diagnoses of HGD and EAC (8).

A field effect has been described in many different cancer types, including in EAC (9, 10). Dysplasia and EAC can be multi-focal in BE. The same mutations and aberrant DNA methylation have been found at multiple levels and in large fields in BE (11, 12), indicating field cancerization. A preneoplastic field surrounding HGD or EAC may appear histologically non-dysplastic (ND) or low grade dysplasia (LGD) but exhibit molecular and cellular changes associated with malignant transformation. Detection of abnormalities in this expanded field may overcome the limitations of random sampling and subjective diagnoses, enabling earlier diagnosis of HGD and EAC.

Many biomarkers have been evaluated in BE (13–16) and the British Society of Gastroenterology (BSG) recommends p53 immunohistochemistry (IHC) to aid diagnosis of dysplasia (17). However, no biomarkers have been validated to reliably detect the field effect or abnormalities associated with prevalent dysplasia and EAC in BE. A tissue systems pathology approach based upon an imaging platform that quantifies both epithelial and stromal abnormalities has been shown to aid in distinguishing HGD from non-dysplastic BE with reactive atypia (18, 19). This imaging approach has also been demonstrated to predict incident progression in BE, by objectively quantifying molecular and cellular features that precede definitive morphologic changes (20) (Figure 1). The assay employs multiplexed

immunofluorescence labeling of 9 epithelial and stromal biomarkers in sections from formalin-fixed paraffin-embedded (FFPE) biopsies. The fluorescently-labeled slides are imaged by whole slide fluorescence scanning, and automated image analysis software extracts quantitative expression and localization data on the biomarkers and morphology. The final step utilizes a multivariable classifier to integrate the quantitative image analysis data into individualized scores that are correlated with risk of HGD/EAC (Figure 1 (20)). This may have applications in detecting molecular and cellular changes in the expanded preneoplastic field associated with HGD/EAC. The aim of this study was to determine whether this assay can detect abnormalities indicative of a field effect in ND, indefinite for dysplasia (IND) and LGD biopsies from BE patients with prevalent HGD/EAC.

Materials and Methods

Study Design and Patients

A case-control study was constructed that utilized a multi-center cohort of BE patients with clinical outcome data from four high volume institutions (Geisinger Health System, University of Pittsburgh, University of Pennsylvania and Academic Medical Center (AMC), Amsterdam, Netherlands). BE cases with ND, IND or LGD confirmed by a gastrointestinal (GI) subspecialist pathologist (J.M.D., J.L., N.C.J.) were retrieved. For patients with multiple biopsy levels taken at the same endoscopy, the biopsy with the highest diagnosis determined by a GI subspecialist pathologist was selected (LGD was the highest diagnosis, then IND, and ND was the lowest). For patients with multiple biopsy levels with the same diagnosis, the pathologist at each institution selected a representative block with sufficient tissue for analysis. Inclusion criteria were availability of tissue blocks and clinicopathologic data, and confirmation of intestinal metaplasia by a GI subspecialist. Exclusion criteria were insufficient tissue quality (assessed by a pathologist), and use of Bouin's fixative or methylene blue in sample processing that can interfere with fluorescence immunolabeling. Cases were patients who had HGD/EAC on repeat endoscopy in <1 year (n=23) or had prior history of treated HGD/EAC, returned to ND, IND or LGD and had HGD/EAC on repeat endoscopy (n=7) (prevalent cases, n=30 in total). Prevalent cases with and without a prior history as described above were included since both subsets of patients can harbor HGD or early EAC that can be challenging to recognize during endoscopy. The non-progressor controls did not show HGD/EAC on repeat endoscopy and had median HGD/EAC-free surveillance time of 5.6 years (n=145). Data elements collected were: case collection date, original pathologic diagnosis and GI subspecialist diagnosis for the case tested in this study, date and original diagnosis of every surveillance biopsy, progression endpoint (HGD/EAC), HGD/EAC-free surveillance time (time between case tested and HGD/EAC diagnosis or last follow-up), age, sex, and segment length (cm) and segment class (short ≤3cm, long >3cm). The study was approved by the institutional review boards at each institution.

Fluorescence Immunolabeling

5μm sections of FFPE BE biopsies were stained with H&E by standard histology methods. K20, p16INK4a, AMACR, p53, HER2/neu, CD68, COX-2, HIF-1α, and CD45RO were labeled by multiplexed immunofluorescence according to previously described methods (19). The biomarkers were multiplexed in sub-panels of 3 primary antibodies per slide

detected via Alexa Fluor-488, -555 and -647-conjugated secondary antibodies and Hoechst-33342 to label DNA (Life Technologies, Carlsbad, CA).

Whole Slide Imaging

H&E-stained slides were imaged at 20× magnification on a NanoZoomer Digital Pathology scanner (Hamamatsu Photonics, K.K., Japan). Fluorescently-immunolabeled slides were imaged using a standard operating procedure at 20× magnification on a ScanScope FL (Leica BioSystems, Vista, CA) as previously described (19).

Image Analysis

Whole slide fluorescence images were analyzed using the TissueCypher™ Image Analysis Platform (Cernostics, Inc., Pittsburgh, PA), which utilizes automated tissue image analysis algorithms for segmenting cell-based objects and tissue structures (e.g. epithelial and stromal compartments) to allow contextual, quantitative biomarker and morphology feature data collection. The image analysis algorithms have been described in detail previously (19) and are summarized in Supplementary Figure S1. The 15 features employed by the risk classifier (Supplementary Table 1 (20)) were extracted from the fluorescence whole slide tissue images.

Statistical Analyses

A risk prediction classifier was developed in a previous study for prediction of incident progression to HGD/EAC (Figure 1 (20)). In this study, we tested the hypothesis that the patients in the predicted high-risk class have significantly higher risk for presence of prevalent HGD/EAC than patients in the predicted low-risk class. We also tested the hypothesis that the risk classes would provide independent and stronger prognostic information beyond that of the pathologic diagnosis (GI subspecialist or generalist pathologist). Sample size calculations indicated that a total of 43 patients (including both prevalent cases and non-progressors) were required to ensure 80% power to detect a significant difference of 50% in the risk of prevalent HGD/EAC between those classified as high-risk vs low-risk, at a 0.05 significance level. All assay parameters were pre-specified, including the 15 image analysis feature/measures, scaling parameters, the classifier model and cutoffs as defined in the previous study (20). The assay parameters are summarized in Figure 1 and Supplementary Table S1). The risk score and risk class (low, intermediate or high) were calculated for each case.

Receiver Operating Characteristic (ROC) curves were plotted based on the binary outcome of the subsequent diagnosis of HGD/EAC (cases) versus no disease progression (control) and the continuous risk scores of the test. ROC curves were also plotted for percentage of cells overexpressing p53 (determined by the image analysis software as described previously (19)). The comparison to p53 was done since the BSG recommends p53 IHC to aid in the diagnosis of dysplasia (17). Logistic regression was used to evaluate the significance of association of the predicted risk groups as the independent variable with subsequent diagnosis of prevalent HGD/EAC or not as the dependent variable. Odds ratios (ORs) with 95% C.I. measuring the strength of the association between the predicted risk groups and the subsequent diagnosis of HGD/EAC were calculated from the logistic regression.

Comparison of Classifier Performance versus Pathologic Diagnosis

Multivariate logistic regression and multivariate Cox regression were performed to compare the performance of the risk classes produced by the classifier versus the pathologic diagnosis by either generalist or GI subspecialist included as the independent variable, in predicting subsequent diagnosis of HGD/EAC included as the dependent variable. The pathologic diagnosis was dichotomized: LGD versus ND and IND combined. IND and ND cases were combined due to the limited sample size of the IND subset (2 non-progressors and 1 prevalent case had subspecialist diagnosis of IND).

Results

Patients

The case-control cohort included biopsies with diagnoses of ND (n=13), IND (n=1) or LGD (n=16) from 30 BE patients with HGD or EAC (prevalent cases, median time to HGD/EAC diagnosis 140.5 days, IQR 56, 241) and 145 samples from matched control patients with clinical outcome data showing no disease progression (ND n=138, IND n=2, LGD n=5, median HGD/EAC-free surveillance time 2,015 days, IQR 1,498, 3,111). 22/30 prevalent cases were diagnosed with HGD and 8/30 were diagnosed with EAC on repeat endoscopy (Table 1). The control patients were from a cohort evaluated in a previous study (20), whereas the prevalent cases had not previously been evaluated. The majority of the patients were male, and a higher proportion of patients with HGD/EAC were male (93.3%) and had long segment BE (63.3%) compared to the non-progressors (78.6% were male, 50.3% had long segment), which is consistent with published epidemiology studies in EAC (21, 22). The clinical characteristics of the patients are summarized in Table 1.

Performance of 15-Feature Risk Classifier in Stratifying Prevalent Cases from Non-Progressor Patients

The pre-specified 15-feature risk classifier was evaluated in the set of BE biopsies from prevalent cases and non-progressor patients. ROC analysis based on the binary outcome (subsequent diagnosis of HGD/EAC versus no disease progression) and the continuous risk scores showed that the classifier had the capability to distinguish prevalent HGD/EAC from non-progressors with AUROC of 0.893, whereas the % cells overexpressing p53 had AUROC 0.594 (Figure 2A). Sub-analyses were performed on the subsets of prevalent cases with and without a prior history of HGD or EAC, since cases with a prior history of HGD/EAC may harbor greater numbers of mutations and other abnormalities than cases with no prior history. ROC analysis of the subsets of prevalent cases with and without a prior history of HGD or EAC (n=7 and 23, respectively) showed that the classifier had strong predictive performance in distinguishing both subsets from non-progressors (AUROC=0.926 and 0.883, respectively). AUROC for the ND/IND and LGD subclasses were 0.873 and 0.792, respectively. A box and whisker plot showed higher 15-feature risk scores in the prevalent cases versus non-progressors ($p<0.0001$, Figure 2B). Logistic regression demonstrated that the 15-feature classifier could stratify patients with significantly different risks for prevalent HGD/EAC; ORs were 46.0 (95% C.I. 14.86–169, $p<0.0001$) for the comparison of the high-risk versus low-risk group and 7.67 (95% C.I. 2.24–28.14, $p=0.001$) for intermediate-risk versus low-risk group (Figure 2C). The classifier identified both non-

dysplastic and LGD biopsies from prevalent cases as high-risk (Figure 2D). The probability of diagnosis of HGD/EAC on repeat endoscopy increased continuously as the 15-feature risk score increased (Figure 2E). In multivariate logistic regression in which subsequent diagnosis of HGD/EAC was evaluated first in relation to pathologic diagnosis alone, then in relation to the predicted risk classes added to the pathologic diagnosis, the magnitude of ORs indicated that the predicted high-risk class provided independent and stronger predictive power than the generalist or GI subspecialist pathologic diagnosis in this cohort of patients (Table 2). Similar results were obtained with multivariate Cox regression models (Supplementary Table S2).

The patients identified as high-risk exhibited multiple epithelial and stromal abnormalities that are quantified by the 15 image analysis features utilized by the risk classifier. Abnormalities detected in ND and LGD biopsies in patients with prevalent HGD/EAC included overexpression of p53, HER2/neu and COX-2, focal AMACR overexpression, infiltration of the lamina propria by CD45RO-positive cells, CD68-positive cells and stromal cells expressing HIF-1α (Figure 3). The stronger predictive power of the risk classes compared to the pathologic diagnosis was illustrated in a patient with 2cm segment BE with biopsies available from two endoscopic levels. Biopsies from 32cm and 34cm were diagnosed as ND and LGD, respectively, by a GI subspecialist in this study, and as ND and IND, respectively, by a general pathologist who recorded the original diagnosis. Repeat biopsy 56 days later showed HGD. Biopsies from the two levels, which were evaluated for illustrative purposes, scored 8.9 and 8.7 (on a scale of 0–10) with the 15-feature risk score, demonstrating that similar high-risk molecular and cellular changes were present at both biopsy levels despite the different pathologic diagnosis (Figure 4A–J).

Discussion

Using a case-control study design we validated a multivariable classifier that assesses ND, IND and LGD biopsies to detect prevalent HGD/EAC in BE patients. The test integrates quantitative biomarker and morphometric data into a risk score, and incorporates 3-tier risk stratification to classify patients as low-, intermediate- or high-risk for HGD/EAC. The predicted high-risk group of patients was at 46-fold increased risk for prevalent HGD/EAC compared to the low-risk group. Importantly, the risk classes provided stronger predictive power than the expert GI and generalist pathologic diagnosis in this cohort of patients, and demonstrated high accuracy in detecting presence of prevalent HGD/EAC, even in non-dysplastic biopsies. The tissue systems pathology assay used in this study thus has the potential to provide physicians and patients with an individualized score that indicates potential for prevalent HGD/EAC, which may aid in decision-making on more rigorous surveillance examinations and endoscopic therapy in BE patients with ND, IND or LGD. This study was strengthened by the use of a diverse patient cohort from four high-volume institutions in the US and Europe. The study was further strengthened by the assay technology, which evaluates multiple pathways associated with carcinogenesis, and is also quantitative and objective. The assay can be performed on sections from FFPE blocks and is thus compatible with clinical practice.

Expert referral centers have higher rates of detection of HGD/EAC and mucosal abnormalities than community centers, but recognition of subtle lesions containing HGD and early EAC can be challenging in all settings (7, 23). Endoscopic surveillance is effective when done in accordance with practice guidelines, however, adherence to the guidelines varies between settings (23, 24). HGD and early EAC may be missed by random sampling, which can result in repeat endoscopies and delays in diagnosis and treatment of HGD and EAC. A diagnosis of LGD confirmed by multiple GI subspecialists is a strong predictor of malignant progression (25, 26). However, intra-observer variability in the diagnosis, even among GI subspecialists, has been well documented (6, 27).

Despite extensive efforts to identify and validate biomarkers in BE (13–16) none have yet been translated into practice to overcome the limitations of random sampling via detection of a field effect. p53 IHC has been demonstrated to have diagnostic and prognostic significance in BE (17, 28). However, assessment of p53 alone is not sufficient since not all patients have detectable abnormalities in p53 protein levels, and a subset of patients who exhibit p53 abnormalities do not develop HGD or EAC (14, 29). Molecular approaches such as DNA sequencing, gene expression, mutation and methylation profiling have been applied to diagnostic and prognostic testing in BE (30–32) but have not yet been evaluated in the detection of abnormalities in the expanded field surrounding HGD and EAC. These technologies have the disadvantage of requiring tissue digestion, resulting in loss of contextual information, such as nuclear morphology, and spatial relationships that are relevant to patient outcomes. Further disadvantages include the requirement for fresh frozen specimens for some of these genomic approaches, which is a logistical problem in clinical practice, and also the need for laser microdissection of tissue areas based on subjective review for some of these approaches.

The risk classifier evaluated in this study identifies patients who have prevalent HGD/EAC, despite receiving a pathologic diagnosis of ND, IND or LGD. The abnormalities quantified by the assay include loss of tumor suppression, loss of cell cycle control, morphologic changes, increased inflammation, stromal angiogenesis, and altered patterns of infiltrating immune cells (20). If validated in additional studies, this finding suggests that objective detection of multiple molecular and cellular abnormalities in the preneoplastic field of BE could overcome some of the limitations of random endoscopic sampling and pathologic diagnosis, and enable earlier detection of HGD/EAC in all practice settings. Additional studies are required to evaluate the performance of the assay in biopsies rom different endoscopic levels. Ex-vivo tests such as this may also complement newer endoscopic techniques such as volumetric laser endomicroscopy (33) by providing objective, quantitative analysis of pathways involved in malignant transformation.

We readily recognize the limitations of this study, which include the retrospective nature of the study, the case-control cohort study design in which the proportion of prevalent cases was not representative of the general population, and the small number of available prevalent cases as reflected in the wide confidence intervals we report. Additional, larger studies will be required to validate our findings. However, large prospective studies are challenging in BE due to the low prevalence of malignant progression. The set of biopsies and patients was heterogeneous; the biopsies had diagnoses of ND, IND and LGD, 22/30 patients had

prevalent HGD while 8/30 had prevalent EAC, and the intervals between the biopsy tested and the repeat endoscopy demonstrating HGD/EAC were variable. Biopsies from a single endoscopy level were evaluated in this study, and additional studies are required with biopsies from multiple endoscopy levels from patients with subtle lesions containing HGD or early EAC. The cohort in this study included patients in surveillance at multiple centers in the US and Europe over a wide timeframe, which prevented standardization of pre-analytic variables. However, the biopsies reflect routine samples requiring accurate risk assessment.

The assay requires instrumentation and software that could not easily be integrated into current pathology laboratories, which are only beginning to adopt digital pathology. The assay has been deployed in a central reference laboratory equipped with the necessary resources, which enables physicians at expert referral centers and community centers to order the assay. Unstained slides can be sent to the reference laboratory, the assay is performed and the laboratory provides a clinical report to the ordering physician and submitting pathologist. The testing process (Supplementary Figure S1) takes approximately 3 business days. The testing approach would initially add to the cost of BE surveillance. However, risk prediction testing has the potential to result in cost savings in high-risk patients by reducing repeat endoscopies and pathologist time required to diagnose HGD/EAC, enabling earlier intervention with endoscopic therapies to reduce EAC incidence and mortality. In low-risk patients there is the potential to lower future costs by extending surveillance intervals (34, 35). While it is feasible to test biopsies from multiple levels in long segment BE, this may add significant cost in a subset of patients, which may outweigh the potential cost savings from reducing unnecessary endoscopies or intervening early to prevent progression.

In summary, the tissue systems pathology assay examined in this study objectively quantifies multiple epithelial and stromal processes that predict prevalent HGD/EAC in BE patients. The assay has the potential to improve upon current histology methods to enable earlier detection of HGD/EAC, which will facilitate earlier, more effective therapeutic interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Georgiann Noll and Matthew Barley (Geisinger) for assistance with specimen retrieval, and D. Lansing Taylor, Ph.D., Janet Warrington, Ph.D. and Aaron DeWard, Ph.D. for many helpful discussions and advice on the research study.

Financial support: This work was supported by a grant from the Pennsylvania Department of Health Cure Program Grant, Research on Cancer Diagnostics or Therapeutics with Commercialization Potential RFA#10-07-03 (R.J. Critchley-Thorne, J.M. Davison, G.W. Falk, J.W. Prichard, Y. Zhang), a Qualifying Therapeutic Discovery Project Grant, Internal Revenue Service/Affordable Care Act 2010, (R.J. Critchley-Thorne), and by Cernostics, Inc. Partial support by the National Cancer Institute of the National Institutes of Health under Award Number R44CA192416 (R.J. Critchley-Thorne, J.M. Davison).

Abbreviations

AMACR	alpha-methylacyl-CoA racemase
AUROC	area under receiver operating characteristic curve
BE	Barrett's esophagus
K20	cytokeratin-20
COX-2	cyclo-oxygenase-2
EAC	esophageal adenocarcinoma
EMR	endoscopic mucosal resection
HGD	high grade dysplasia
HIF-1α	hypoxia-inducible factor 1-alpha subunit
IND	indefinite for dysplasia
LGD	low grade dysplasia
ND	non-dysplastic
OR	odds ratio
RFA	radiofrequency ablation
ROC	receiver operating characteristic

References

1. Cancer Facts & Figures 2016. American Cancer Society.
2. Bulsiewicz WJ, Kim HP, Dellon ES, Cotton CC, Pasricha S, Madanick RD, et al. Safety and efficacy of endoscopic mucosal therapy with radiofrequency ablation for patients with neoplastic Barrett's esophagus. Clin Gastroenterol Hepatol. 2013; 11:636–642. [PubMed: 23103824]
3. Pasricha S, Bulsiewicz WJ, Hathorn KE, Komanduri S, Muthusamy VR, Rothstein RI, et al. Durability and predictors of successful radiofrequency ablation for Barrett's esophagus. Clin Gastroenterol Hepatol. 12:1840–1847. e1.
4. Phoa KN, van Vilsteren FG, Weusten BL, Bisschops R, Schoon EJ, Ragunath K, 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. [PubMed: 24668102]
5. Shaheen NJ, Falk GW, Iyer PG, Gerson LB. ACG Clinical Guideline: Diagnosis and Management of Barrett's Esophagus. Am J Gastroenterol. 2015; 111:30–50. [PubMed: 26526079]
6. Montgomery E, Bronner MP, Goldblum JR, Greenson JK, Haber MM, Hart J, et al. Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. Hum Pathol. 2001; 32:368–378. [PubMed: 11331953]
7. SchöLvinck D, Van Der Meulen K, Bergman JJ, Weusten BL. Detection of Lesions in Dysplastic Barrett's Esophagus: Are Expert Endoscopists Doing a Better Job Than Community Endoscopists? Gastrointestinal Endoscopy. 2015; 81:AB139.
8. Visrodia K, Singh S, Krishnamoorthi R, Ahlquist DA, Wang KK, Iyer PG, et al. Magnitude of Missed Esophageal Adenocarcinoma After Barrett's Esophagus Diagnosis: A Systematic Review and Meta-analysis. Gastroenterology. 150:599–607. e7.

9. Tabor MP, Brakenhoff RH, van Houten VM, Kummer JA, Snel MH, Snijders PJ, et al. Persistence of genetically altered fields in head and neck cancer patients: biological and clinical implications. Clin Cancer Res. 2001; 7:1523–1532. [PubMed: 11410486]
10. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. Cancer Res. 2003; 63:1727–1730. [PubMed: 12702551]
11. Prevo LJ, Sanchez CA, Galipeau PC, Reid BJ. p53-mutant clones and field effects in Barrett's esophagus. Cancer Res. 1999; 59:4784–4787. [PubMed: 10519384]
12. Eads CA, Lord RV, Kurumboor SK, Wickramasinghe K, Skinner ML, Long TI, et al. Fields of aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. Cancer Res. 2000; 60:5021–5026. [PubMed: 11016622]
13. Barrett MT, Sanchez CA, Galipeau PC, Neshat K, Emond M, Reid BJ. Allelic loss of 9p21 and mutation of the CDKN2/p16 gene develop as early lesions during neoplastic progression in Barrett's esophagus. Oncogene. 1996; 13:1867–1873. [PubMed: 8934532]
14. Bird-Lieberman EL, Dunn JM, Coleman HG, Lao-Sirieix P, Oukrif D, Moore CE, et al. Population-based study reveals new risk-stratification biomarker panel for Barrett's esophagus. Gastroenterology. 2012; 143:927–935. e3. [PubMed: 22771507]
15. Kastelein F, Biermann K, Steyerberg EW, Verheij J, Kalisvaart M, Looijenga LH, et al. Aberrant p53 protein expression is associated with an increased risk of neoplastic progression in patients with Barrett's oesophagus. Gut. 2012; 62:1676–1683. [PubMed: 23256952]
16. Kastelein F, Biermann K, Steyerberg EW, Verheij J, Kalisvaart M, Looijenga LH, et al. Value of alpha-methylacyl-CoA racemase immunochemistry for predicting neoplastic progression in Barrett's oesophagus. Histopathology. 2013; 63:630–639. [PubMed: 24004067]
17. Fitzgerald RC, di Pietro M, Ragunath K, Ang Y, Kang JY, Watson P, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. Gut. 2014; 63:7–42. [PubMed: 24165758]
18. Gough, A., Lezon, T., Faeder, J., Chennubhotla, C., Murphy, R., Critchley-Thorne, R., et al. High Content Analysis And Cellular And Tissue Systems Biology: A Bridge Between Cancer Cell Biology And Tissue-Based Diagnostics. In: Mendelsohn, J.Howley, PM.Israel, MA.Gray, JW., Thompson, CB., editors. The Molecular Basis of Cancer 4th Edition. 4th. New York: Elsevier; 2014.
19. Prichard JW, Davison JM, Campbell BB, Repa KA, Reese LM, Nguyen XM, et al. TissueCypher: A Systems Biology Approach to Anatomic Pathology. Journal of Pathology Informatics. 2015; 6:48. [PubMed: 26430536]
20. Critchley-Thorne RJ, Duits LC, Prichard JW, Davison JM, Jobe BA, Campbell BB, et al. A Tissue Systems Pathology Assay for High-Risk Barrett's Esophagus. Cancer Epidemiol Biomarkers Prev. 2016; 25:958–968. [PubMed: 27197290]
21. Mathieu LN, Kanarek NF, Tsai HL, Rudin CM, Brock MV. Age and sex differences in the incidence of esophageal adenocarcinoma: results from the Surveillance, Epidemiology, and End Results (SEER) Registry (1973–2008). Dis Esophagus. 2013; 27:757–763. [PubMed: 24118313]
22. Weston AP, Badr AS, Hassanein RS. Prospective multivariate analysis of clinical, endoscopic, and histological factors predictive of the development of Barrett's multifocal high-grade dysplasia or adenocarcinoma. Am J Gastroenterol. 1999; 94:3413–3419. [PubMed: 10606296]
23. Abrams JA, Kapel RC, Lindberg GM, Saboorian MH, Genta RM, Neugut AI, et al. Adherence to biopsy guidelines for Barrett's esophagus surveillance in the community setting in the United States. Clin Gastroenterol Hepatol. 2009; 7:736–742. quiz 10. [PubMed: 19268726]
24. Verbeek RE, Leenders M, Ten Kate FJ, van Hillegersberg R, Vleggaar FP, van Baal JW, et al. Surveillance of Barrett's esophagus and mortality from esophageal adenocarcinoma: a population-based cohort study. Am J Gastroenterol. 2014; 109:1215–1222. [PubMed: 24980881]
25. Curvers WL, ten Kate FJ, Krishnadath KK, Visser M, Elzer B, Baak LC, et al. Low-grade dysplasia in Barrett's esophagus: overdiagnosed and underestimated. Am J Gastroenterol. 2010; 105:1523–1530. [PubMed: 20461069]

26. Duits LC, Phoa KN, Curvers WL, Ten Kate FJ, Meijer GA, Seldenrijk CA, et al. Barrett's oesophagus patients with low-grade dysplasia can be accurately risk-stratified after histological review by an expert pathology panel. Gut. 2015; 64:700–706. [PubMed: 25034523]
27. Kerkhof M, van Dekken H, Steyerberg EW, Meijer GA, Mulder AH, de Bruine A, et al. Grading of dysplasia in Barrett's oesophagus: substantial interobserver variation between general and gastrointestinal pathologists. Histopathology. 2007; 50:920–927. [PubMed: 17543082]
28. Horvath B, Singh P, Xie H, Thota PN, Sun X, Liu X. Expression of p53 predicts risk of prevalent and incident advanced neoplasia in patients with Barrett's esophagus and epithelial changes indefinite for dysplasia. Gastroenterol Rep (Oxf). 2015
29. Sikkema M, Kerkhof M, Steyerberg EW, Kusters JG, van Strien PM, Looman CW, et al. Aneuploidy and overexpression of Ki67 and p53 as markers for neoplastic progression in Barrett's esophagus: a case-control study. Am J Gastroenterol. 2009; 104:2673–2680. [PubMed: 19638963]
30. Jin Z, Cheng Y, Gu W, Zheng Y, Sato F, Mori Y, et al. A multicenter, double-blinded validation study of methylation biomarkers for progression prediction in Barrett's esophagus. Cancer Res. 2009; 69:4112–4115. [PubMed: 19435894]
31. Eluri S, Brugge WR, Daglilar ES, Jackson SA, Styn MA, Callenberg KM, et al. The Presence of Genetic Mutations at Key Loci Predicts Progression to Esophageal Adenocarcinoma in Barrett's Esophagus. Am J Gastroenterol. 2015; 110:828–834. [PubMed: 26010308]
32. Del Portillo A, Lagana SM, Yao Y, Uehara T, Jhala N, Ganguly T, et al. Evaluation of Mutational Testing of Preneoplastic Barrett's Mucosa by Next-Generation Sequencing of Formalin-Fixed, Paraffin-Embedded Endoscopic Samples for Detection of Concurrent Dysplasia and Adenocarcinoma in Barrett's Esophagus. J Mol Diagn. 2015; 17:412–419. [PubMed: 26068095]
33. Evans JA, Poneros JM, Bouma BE, Bressner J, Halpern EF, Shishkov M, et al. Optical coherence tomography to identify intramucosal carcinoma and high-grade dysplasia in Barrett's esophagus. Clin Gastroenterol Hepatol. 2006; 4:38–43. [PubMed: 16431303]
34. Hao, J., Snyder, SR., Pitcavage, J., Critchley-Thorne, RJ. 21st Annual International Society for Pharmacoeconomics and Outcomes Research Meeting. Washington DC: 2016. A Cost-effectiveness analysis of a test that predicts risk of malignant progression in Barrett's esophagus.
35. Gordon LG, Mayne GC, Hirst NG, Bright T, Whiteman DC, Watson DI. Cost-effectiveness of endoscopic surveillance of non-dysplastic Barrett's esophagus. Gastrointest Endosc. 2014; 79:242–256. e6. [PubMed: 24079411]

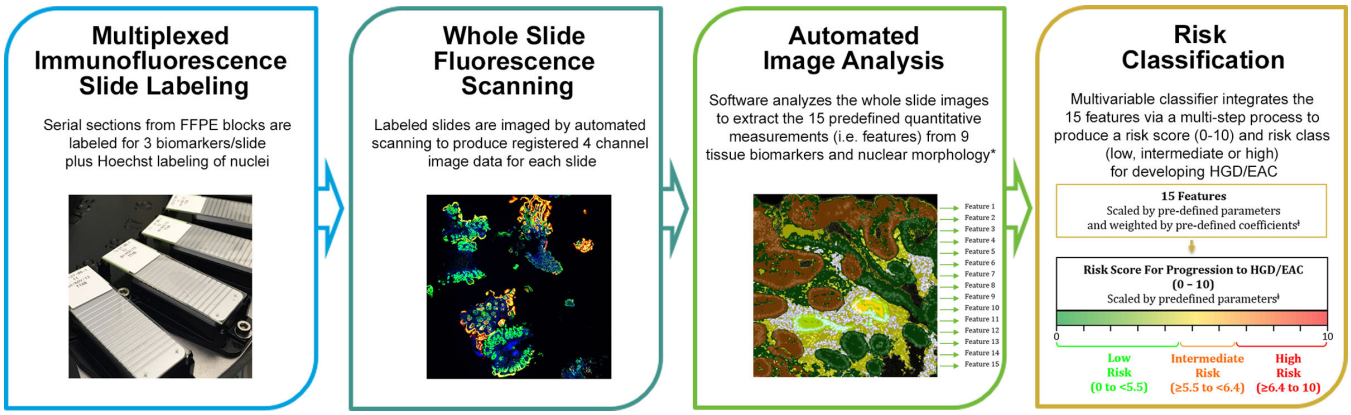
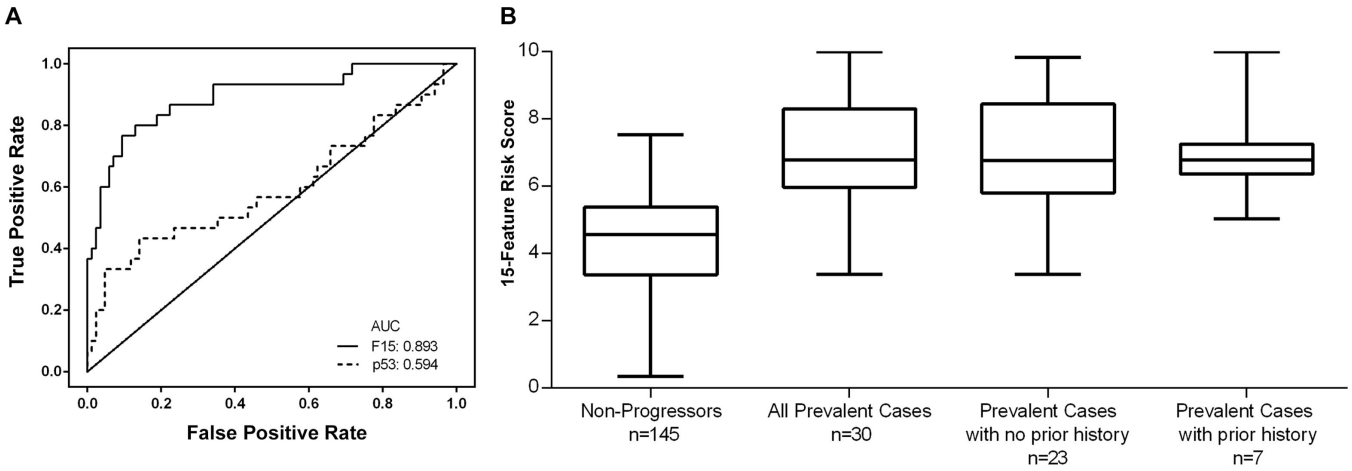


Figure 1. 15-Feature 3-Tier Risk Classifier Process

The risk score (0–10) and class (low, intermediate or high) are calculated from the scaled and coefficient-weighted sum of 15 quantitative image analysis measurements (features) derived from 9 protein-based biomarkers and morphology as follows: *1) Multiplexed Immunofluorescence Slide Labeling* - Serial sections of FFPE BE biopsies are fluorescently immunolabeled for p16, AMACR, p53, HER2, K20, CD68, COX-2, HIF-1α, and CD45RO, plus Hoechst; *2) Whole Slide Fluorescence Scanning* - Labeled slides are imaged by whole slide fluorescence scanning that generates image data on each biomarker and nuclei; *3) Automated Image Analysis*: Tissue images are analyzed by automated image analysis software to extract 15 features rom the 9 protein-based biomarkers and Hoechst, (Supplementary Table S1); *4) Risk Classification*: The 15 features are scaled using center and scale parameters defined in a training study (20), then weighted by coefficients derived from univariate Cox regression analysis of the features and progression outcomes in the training study (20). The weighted sum of the 15 scaled features produces an unscaled risk score, which is scaled as follows:

$$\text{Risk Score} = \begin{cases} 0 & \text{if raw score} < -10 \\ \frac{\text{raw score} + 10}{2} & \text{if } -10 < \text{raw score} < 10 \\ 10 & \text{if raw score} > 10 \end{cases}$$

Cutoffs are applied to the risk score to classify patients for risk of progression: Risk Class=Low if scaled score 0–<5.5, =Intermediate if scaled score ≥5.5<6.4, =High if scaled score ≥6.4–10.



C

Comparison	OR (95% C.I.)	P value
Inter Risk vs. Low Risk	7.67 (2.24 - 28.14)	0.001
High Risk vs. Low Risk	46.0 (14.86 - 169)	< 0.0001

D

Patient Subset	Pathologic Diagnosis	Low Risk	Inter Risk	High Risk
Prevalent Cases	ND/IND	3	4	8
	LGD	2	3	10
Non-Progressors	ND/IND	114	19	7
	LGD	1	2	2

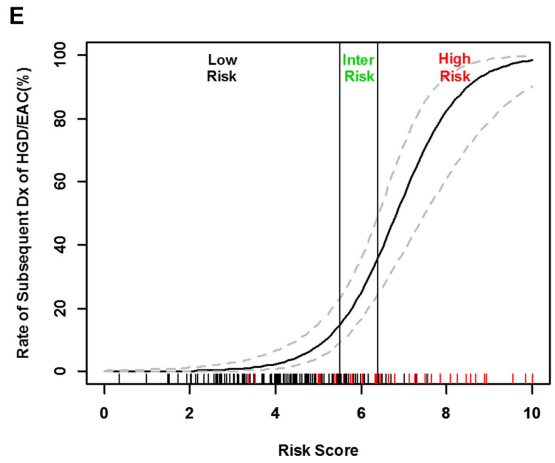


Figure 2. Performance of 15-Feature Risk Score in Non-Dysplastic and LGD BE Biopsies from Non-Progressor Patients and Patients with Prevalent HGD/EAC

A: ROC curve for 15-feature risk score and percentage of cells overexpressing p53 (determined by image analysis software as described previously (19). **B:** Box and whisker plots of the 15-feature risk score in non-progressors and prevalent cases ($p<0.0001$, Wilcoxon rank sum test comparing non-progressors vs. all prevalent cases). **C:** Univariate ORs with 95% C.I. and p-values from logistic regression for comparisons between the predicted risk classes. **D:** Number of cases scored low-, intermediate (inter)-, and high-risk by GI subspecialist pathologic diagnosis. **E:** Rate of subsequent diagnosis of HGD/EAC as a continuous function of the 15-feature risk score. Dashed curves indicate 95% C.I. The rug plot on the x-axis shows the risk score for non-progressor controls (black dashes) and prevalent cases (red dashes), andcutoffs for low-, inter-, and high-risk are shown.

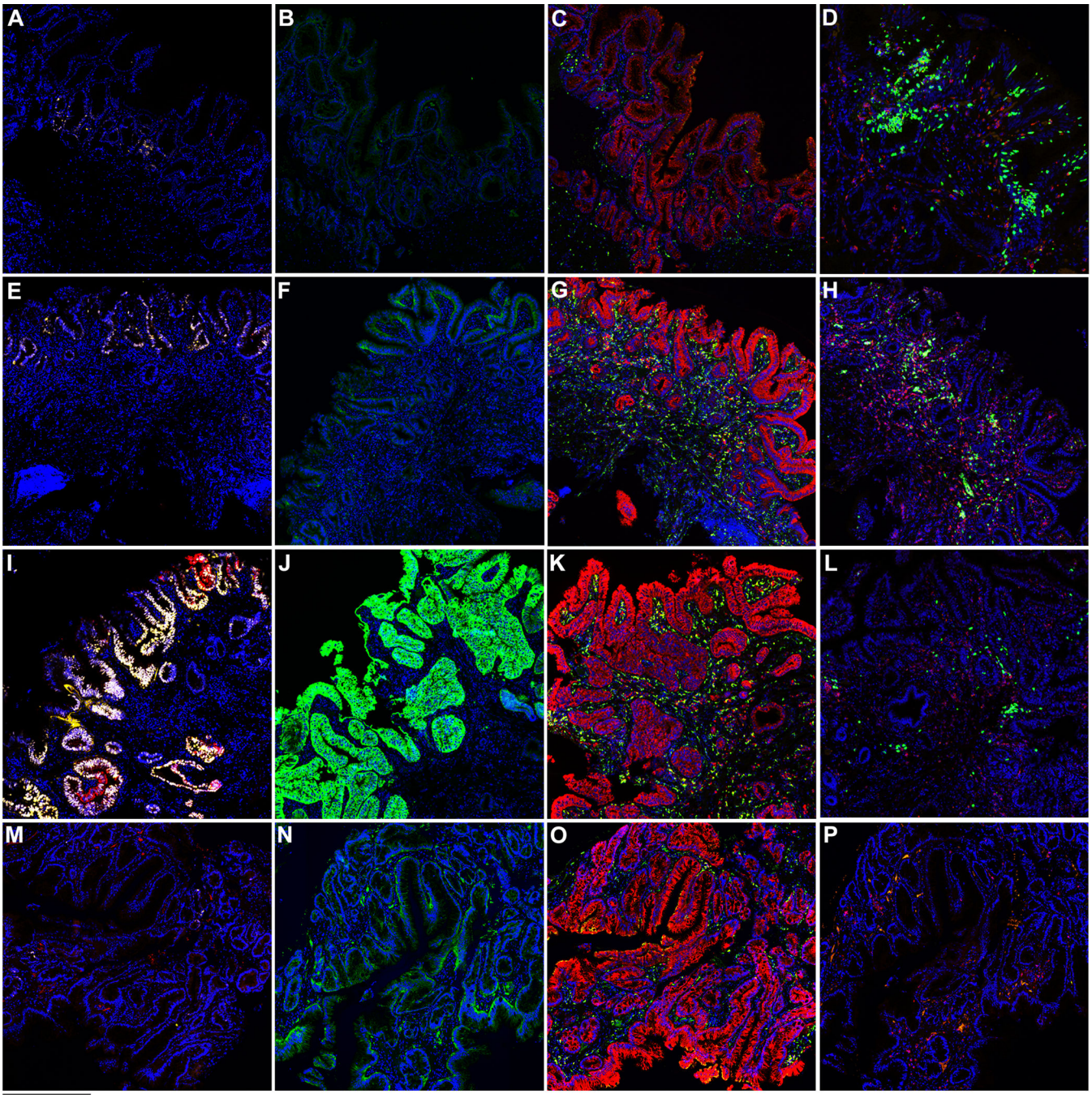


Figure 3. Representative Images of High Risk Biomarkers in BE Biopsies

Panels A–D show an ND biopsy from a patient who had HGD on repeat endoscopy 310 days later; **A:** p53-yellow, AMACR-red, **B:** HER2/neu-green. **C:** CD68-green, COX-2-red, **D:** HIF-1α-green, CD45RO-red. Panels E–H show a LGD biopsy from a patient who had HGD on repeat endoscopy 56 days later; **E:** p53-yellow, AMACR-red, **F:** HER2/neu-green. **G:** CD68-green, COX-2-red, **H:** HIF-1α-green, CD45RO-red. Panels I–L show a LGD biopsy from a patient who had HGD on repeat endoscopy 60 days later; **I:** p53-yellow, AMACR-red, **J:** HER2/neu-green. **K:** CD68-green, COX-2-red, **L:** HIF-1α-green, CD45RO-red.

Panels M–P show a ND biopsy from a non-progressor patient with HGD/EAC-free surveillance time of 2,186 days; **M:** p53-yellow, AMACR-red, **N:** HER2/neu-green. **O:** CD68-green, COX-2-red, **P:** HIF-1α-green, CD45RO-red. Hoechst labeling of nuclei is shown in blue in all panels.

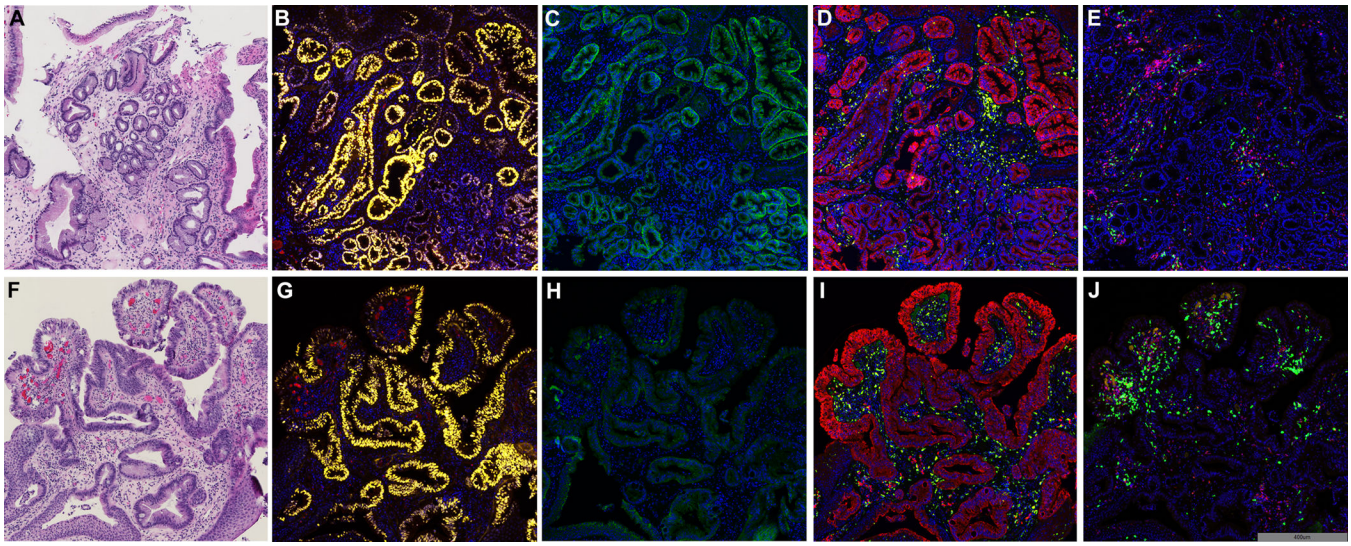


Figure 4. Representative Images of High Risk Biomarkers and Risk Scores at Multiple Endoscopic Levels

Panels A–E and Panels F–J show a LGD biopsy and a ND biopsy, respectively, from a patient with 2cm segment BE who had HGD on repeat endoscopy 56 days later. The images show similar epithelial and stromal abnormalities in the biopsies despite the difference in diagnosis. The 15-feature risk scores for the LGD and ND biopsies were 8.7 and 8.9 (both high-risk), respectively. Panels A–E LGD biopsy - **A:** H&E, **B:** p53-yellow, AMACR-red, **C:** HER2/neu-green. **D:** CD68-green, COX-2-red, **E:** HIF-1α-green, CD45RO-red. Panels F–J ND biopsy - **F:** H&E, **G:** p53-yellow, AMACR-red, **H:** HER2/neu-green. **I:** CD68-green, COX-2-red, **J:** HIF-1α-green, CD45RO-red. Hoechst labeling of nuclei shown in blue.

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2018 February 01.

Table 1

Patient Characteristics

	Non-Progressors				Prevalent Cases			
# patients	145				30			
HGD/EAC-free surveillance time (Median days (IQR)	2,015 (1,498, 3,111)				140.5 (56, 241.3)			
Age (mean years ±S.D.)	61.0 ±12.1				61.8 ±9.5			
Segment Length (%)								
Short (≤3cm)	58 (40.0)				9 (30.0)			
Long (>3cm)	73 (50.3)				19 (63.3)			
Unknown	14 (9.7)				2 (6.7)			
Sex (%)								
Male	114 (78.6)				28 (93.3)			
Female	31 (21.4)				2 (6.7)			
Patients in each diagnostic class (%)	ND	IND		LGD	ND	IND		LGD
	138 (95.2)	2 (1.4)		5 (3.4)	13 (43.3)	1 (3.3)		16 (53.3)
End point diagnoses (%)	na	na		na	HGD 7 (53.8) EAC 6 (46.2)	HGD 1 (100) EAC 0 (0)		HGD 14 (87.5) EAC 2 (12.5)
Each Institution	**Non-Progressors**				**Prevalent Cases**			
# patients (%)	AMC	Gei	UPenn	UPitt	AMC	Gei	UPenn	UPitt
	46 (31.7)	63 (43.5)	15 (10.3)	21 (14.5)	4 (13.3)	12 (40.0)	7 (23.3)	7 (23.3)
HGD/EAC-free surveillance time (median days (IQR)	1,816 (1,458, 2,847.3)	2,361 (1,522, 3,698)	1,497 (1,277, 1,736)	2,306 (2,055, 3,155.5)	98.5 (60.5, 172.5)	42 (19, 134.5)	220 (176, 321)	179 (123, 242)

*surveillance time: number of days between biopsy tested and last endoscopy with ND, IND or LGD (non-progressors) or endoscopy with diagnosis HGD or EAC (prevalent cases). Diagnosis provided by GI subspecialist pathologist.

AMC: Academic Medical Center, Netherlands; Gei: Geisinger Health System; UPenn: University of Pennsylvania; UPitt, University of Pittsburgh. S.D.: standard deviation, IQR: interquartile range.

Table 2

Performance of Risk Classes Predicted by Test vs. Pathologic Diagnosis in Stratifying BE Patients with Prevalent HGD/EAC from Non-Progressor BE Patients.

A. Predictive Performance of Risk Classes vs. Generalist Pathologist Diagnosis		
Variable	**Odds Ratio (95% CI)**	**P Value**
Analysis *without* Risk Prediction Test		
General Pathologist's Dx (LGD vs. ND/IND)	12.67 (4.17 – 44.05)	<0.0001
Analysis *with* Risk Prediction Test		
General Pathologist's Dx (LGD vs. ND/IND*)*	5.28 (1.42 – 21.39)	0.01
Risk Classes (predicted by the test)		
Intermediate vs. Low Risk	12.23 (2.19 – 95.92)	0.007
High vs. Low Risk	32.16 (6.40 – 246.94)	0.0001
B. Predictive Performance of Risk Classes vs. GI Subspecialist Pathologist colspan="3" Diagnosis		
Variable	**Odds Ratio (95% CI)**	**P Value**
Analysis *without* Risk Prediction Test		
GI Subspecialist Pathologist's Dx (LGD vs. ND/IND)	28.0 (9.47 – 96.63)	<0.0001
Analysis *with* Risk Prediction Test		
GI Subspecialist Pathologist's Dx (LGD vs. ND/IND)	10.36 (2.85 – 42.46)	0.0006
Risk Classes (predicted by the test)		
Intermediate vs. Low Risk	5.16 (1.34 – 20.34)	0.01
High vs. Low Risk	24.65 (7.15 – 96.58)	<0.0001

Multivariate logistic regressions were run in which subsequent diagnosis of HGD/EAC as the dependent variable was evaluated first in relation to pathologic diagnosis alone, then in relation to risk classes and pathologic diagnosis, included as the independent variable in non-progressors and prevalent cases. Variables were dichotomized; diagnosis: LGD vs. ND/IND combined, predicted risk classes: intermediate- vs. low-risk class and high- vs. low-risk class. Part A, n=130 patients with generalist diagnosis recorded during surveillance. Part B, n=175 patients with GI subspecialist diagnosis provided for this study.