A LIQUID BIOPSY TO IDENTIFY BARRETT'S ESOPHAGUS, DYSPLASIA, AND ESOPHAGEAL ADENOCARCINOMA: THE *EMERALD* MULTI-CENTER STUDY.

Jinsei Miyoshi[1,2,3] *, Alessandro Mannucci[4,5] *, Marco Scarpa[6], Feng Gao[7], Shusuke Toden[1], Timothy Whitsett[8], Landon J. Inge[9], Ross M. Bremner[9], Tetsuji Takayama[2], Yulan Cheng[10], Teodoro Bottiglieri[11], Iris D Nategaal[12], Martha J. Shrubsole[13], Ali H. Zaidi[14], Xin Wang[15,16,17], Helen G Coleman[18], Lesley A. Anderson[19], Stephen J. Meltzer[10], and Ajay Goel[1,4,20], on behalf of the FINBAR-EMERALD collaborative group

1. Center for Gastrointestinal Research; Center from Translational Genomics and Oncology, Baylor Scott & White Research Institute and Charles A. Sammons Cancer Center, Baylor University Medical Center, Dallas, TX, USA
2. Department of Gastroenterology and Oncology, Tokushima University Graduate School of Biomedical Sciences, Tokushima, Japan
3. Department of Gastroenterology, Kawashima Hospital, Tokushima, Japan.
4. Department of Molecular Diagnostics and Experimental Therapeutics, Beckman Research Institute of City of Hope, Monrovia, California, USA
5. Gastroenterology and Gastrointestinal Endoscopy Unit, Vita-Salute San Raffaele University, IRCCS San Raffaele Hospital, Vita-Salute San Raffaele University, Milan, Italy
6. Department of Surgical, Oncological and Gastroenterological Sciences, University of Padova, Padova, Italy
7. The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
8. Cancer and Cell Biology Division, The Translational Genomics Research Institute (TGen), Phoenix, AZ, USA
9. Norton Thoracic Institute, St Joseph's Hospital and Medical Center, Phoenix, Arizona
10. Division of Gastroenterology and Hepatology, Department Of Medicine And Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, MD, USA
11. Baylor Scott & White Research Institute, Institute of Metabolic Diseases, Dallas, TX, USA
12. Radboud University Medical Centre, Department of Pathology, Nijmegen, the Netherlands
13. Department of Medicine, Division of Epidemiology, Vanderbilt University School of Medicine, Nashville, TN, USA
14. Allegheny Health Network Cancer Institute, Allegheny Health Network, Pittsburgh, PA, USA
15. Department of Surgery, The Chinese University of Hong Kong, Hong Kong, China
16. Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, China
17. Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China
18. Cancer Epidemiology Research Group, Centre for Public Health, Queen's University Belfast, Belfast, United Kingdom
19. Centre for Health Data Science, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK

20. City of Hope Comprehensive Cancer Center, Duarte, California, USA

** Denotes shared first authorship*

Running title: A blood test for Barrett's esophagus and adenocarcinoma

Corresponding Author: Prof. Ajay Goel, PhD, AGAF
Department of Molecular Diagnostics and Experimental Therapeutics, Beckman Research Institute of City of Hope, Duarte, California, USA.
Phone: 626-218-0541
AJGOEL@COH.ORG

Keywords:
Esophagus adenocarcinoma; Barrett's esophagus; High-grade dysplasia; Biomarker; microRNA

Funding: The present work was supported by the CA72851, CA181572, CA184792, CA187956, and CA202797 grants from the National Cancer Institute, National Institute of Health, and a grant from Guangdong Basic and Applied Basic Research Foundation (Project No. 2019B030302012), Shenzhen Medical Research Fund (Project No. C2303002), and grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. 14111522, C4024-22GF, R4007-23) awarded to Xin Wang. This work was also partially sponsored by Shenzhen Bay Scholars Program awarded to Xin Wang. The FINBAR study was funded by Cancer Focus Northern Ireland (formerly the Ulster Cancer Foundation), the Northern Ireland R&D office, and the Health Research Board, Ireland. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosures:
The authors have no conflicts of interest to disclose.

Acknowledgments:

The FINBAR-EMERALD collaborators included Matteo Fassan and Elisabetta Trevellin (both from the Department of Medicine, University of Padova, Italy) and Melania Scarpa (Lab of Advanced Translational Research, Veneto Institute of Oncology, IOV-IRCCS, Padova, Italy). The Factors INfluencing the Barrett's Adenocarcinoma Relationship (FINBAR) study group members included the late Professor Liam J. Murray (Queen's University Belfast), L. A. A. (Queen's University Belfast), B. T. Johnston (Belfast Health & Social Care Trust), R. G. P. Watson (Belfast Health & Social Care Trust), J. McGuigan (Belfast Health & Social Care Trust), H. R. Ferguson (Belfast Health & Social Care Trust), S. J. Murphy (St Vincent's Hospital Dublin), J. V. Reynolds (St James' Hospital, Dublin) and H. Comber (National Cancer Registry of Ireland). Additionally, the authors thank Drs Caiming Xu, Silei Sui, Takayuki Noma, and Yuan Li for their thoughtful discussions and advice during this project.

Patient and Public Involvement statement:

Although the patients were not directly involved in formulating the research question(s) and outcome measures or in designing the study, the authors wish to express their appreciation for the contributions made by the study participants, their families, and all those who assisted with the study, particularly the Northern Ireland Cancer Registry and National Cancer Registry Cork.

Author contributions (CRediT):

Drs. Miyoshi and Mannucci had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conceptualization	JM, AM, ST, XW, AG
Data curation	JM, AM, FG, ST, AG
Formal analysis	JM, AM, FG, AG
Funding acquisition	XW, AG
Investigation	JM, AM, FG, AG
Methodology	JM, AM, FG, ST, XW, AG
Project administration	MS, TW, YC, TB, IDN, HGC, XW, AG
Resources	MS, TW, LI, YC, TB, IDN, HGC, XW, AG
Software	JM, AM, ST, FG, XW, AG
Supervision	MS, LI, TT, AHZ, SJM, HGC, XW, AG
Validation	JM, AM, MS, MJS, AHZ, SJM, HGC, LAA, AG
Visualization	AM, AG
Writing – original draft	AM, AG
Writing – review/edits	JM, AM, FG, ST, SM, TW, LI, TT, YC, TB, IDN, MJS, AHZ, SJM, HGC, LAA, XW, AG.

Manuscript word count: 4425/**4000**

Abstract word count: 259/**250**

Main tables and figures: 8 (1 table and 7 figures)

Supplementary tables and figures: **11** (3 tables and 8 figures)

Number of References: 50/**50**

Additional information:

- **Study registration number:** **NCT06381583** (ClinicalTrial.gov ID)
- **Protocol access:** "Protocol" document, attached to the submission, and at *https://clinicaltrials.gov*
- **Research checklist:** **STARD** (Supplementary Material) and **TRIPOD-AI** (Supplementary Material)
- **Data and code sharing information:** Data collected for the study, including de-identified participant data, and the code will be made available to others at publication via a signed data access agreement and at the discretion of the investigators' approval of the proposed use of such data

ABSTRACT *(259 words)*

Background: There is no clinically relevant serological marker for the early detection of esophageal adenocarcinoma (EAC) and its precursor lesion, Barrett's esophagus (BE).

Objective: To develop and test a blood-based assay for EAC and BE

Design: EMERALD was a large, international, multi-center biomarker cohort study involving 792 patient samples from four countries (NCT06381583) to develop and validate a circulating miRNA signature for the early detection of EAC and high-risk BE. Tissue-based miRNA sequencing and microarray datasets (n=134) were used to identify candidate miRNAs of diagnostic potential, followed by validation using 42 pairs of matched cancer and normal tissues. The usefulness of the candidate miRNAs was initially assessed using 108 sera (44 EAC, 34 EAC precursors, and 30 non-disease controls). We finally trained a machine learning model (XGBoost+AdaBoost) on RT-qPCR results from circulating miRNAs from a training cohort (n=160) and independently tested it in an external cohort (n=295).

Results: After a strict process of biomarker discovery and selection, we identified 6 miRNAs that were over-expressed in all sera of patients compared to non-disease controls from three independent cohorts of different nationalities (miR-106b, -146a, -15a, -18a, -21, and -93). We established a 6-miRNA diagnostic signature using the training cohort (AUROC: 97.6%) and tested it in an independent cohort (AUROC: 91.9%). This assay could also identify patients with BE among patients with gastroesophageal reflux disease (AUROC: 94.8%, sensitivity: 92.8%, specificity: 85.1%).

Conclusion: Using a comprehensive approach integrating unbiased genome-wide biomarker discovery and several independent experimental validations, we have developed and validated a novel blood test that might complement screening options for BE/EAC.

KEY MESSAGES

What is already known on this topic: There are no non-invasive biomarkers for early detection of esophageal adenocarcinoma (EAC) and its precursor lesions, the second most lethal gastrointestinal malignancy. A blood-based test would complement the screening options available and likely improve patient outcomes.

What this study adds: We have developed and independently validated the robustness of a blood-based test named '*EMERALD'* for the early detection of both EAC and Barrett's esophagus precancerous lesions in the largest multi-centric biomarker cohort to date, an effort that involved institutes from four countries.

How this study might affect research, practice, or policy: The *EMERALD* blood test could enrich the toolkit of methods available for EAC screening. Successful implementation of the *EMERALD* assay would be expected to improve patient outcomes. Although the model may support a more cost-effective approach with five-yearly *EMERALD* screening, a head-to-head study would provide definitive evidence to guide clinical practice.

INTRODUCTION

Esophageal adenocarcinoma (EAC), the second most lethal gastrointestinal malignancy after pancreatic cancer, is linked to chronic exposure to gastroesophageal reflux disease (GERD) and the development of Barrett's esophagus (BE) [1-3]. The progression from BE to low-grade dysplasia (LGD), high-grade dysplasia (HGD), and ultimately EAC unfolds slowly over approximately 20 years (1-3% risk year) [4-6]. This extended timeframe theoretically offers ample opportunities for cancer prevention. However, fewer than 20% of patients with BE receive a diagnosis before they are diagnosed with EAC [7-9]. Therefore, most EAC cases are diagnosed *de novo*, bypassing the window for preventive interventions [10,11]. Moreover, the lethality of EAC is further explained by its rapid progression from a localized stage to regional and distant metastases because the esophageal anatomy, devoid of a serosa but rich in a dense lymphatic network, offers minimal resistance against the rapid and early spread of cancer [12].

EAC represents a public health threat: its incidence has risen dramatically since the 1980s, and despite advances in therapy, the overall 5-year survival has remained below 20% [13-15]. These observations, coupled with evidence for the cost-effectiveness of endoscopic screening [16-20], have led to recommendations for endoscopy for patients with persistent GERD or risk factors for BE and EAC [21-25]. While early detection of EAC alone would reduce mortality but not incidence, early detection of BE, followed by BE surveillance and treatment for LGD/HGD, can halt the progression to EAC with low rates of recurrence [26-28]. However, the current reliance on endoscopy has limitations that include its invasive nature, costs, and potential discomfort, contributing to poor patient adherence to screening programs [29]. Given the prevalence of GERD and the growing concern over EAC, new clinical strategies complementing current guidelines could be highly beneficial. A minimally invasive approach, such as a liquid biopsy targeting both pre-cancerous lesions and early-stage EAC, may improve patient compliance. This study aimed to address this need by developing a diagnostic model of EAC and its precursor lesions, leveraging state-of-the-art machine learning (ML) driven by biological and clinical data.

ML involves identifying patterns within data and fitting models to the endpoint of interest [30]. Ensemble classifiers combine multiple weak learners (*i.e.*, the boosting procedure) to achieve higher accuracy [30]. To enhance the accuracy of these models, stacking involves creating a "meta-model" on the predictions of the base-line models [30]. Importantly, models like XGBoost and AdaBoost prioritize accuracy alongside interpretability, enabling analysis of biomarker importance with techniques like SHAP values analysis [30].

MicroRNAs (miRNAs), non-coding single-stranded RNAs regulating gene expression and various cellular processes, have been involved in EAC pathogenesis[31]. Due to their stability in body fluids and disease specificity, circulating miRNAs are potentially promising candidates for developing non-invasive liquid biopsies [32]. Individual biomarkers alone are often not discriminative enough for cancer detection. Therefore, biomarker panels are created to combine multiple biomarkers and increase the test's performance by virtue of ML approaches. Previous reports described potential circulating miRNAs for the diagnosis of EAC, but these studies lacked a systematic and comprehensive biomarker discovery approach or have not validated these biomarkers in multiple independent patient cohorts with adequate statistical power [33,34]. In addition, most of these studies have focused essentially on a single or handful of biomarkers, which has resulted in limited sensitivities and specificities [33,34].

In this research effort, we leveraged ML and circulating miRNAs to develop a liquid biopsy assay diagnostic for EAC and BE. By integrating a systematic genome-wide biomarker discovery and clinical validation approach in more than 750 tissue and blood specimens from multiple independent patient cohorts with EAC, HGD, LGD, BE, and healthy subjects from five countries (USA, UK, Ireland, Italy, Netherlands), we have identified, developed, and established a novel liquid biopsy assay ("*EMERALD*" – **E**sophageal **M**icroRNAs of Ba**RR**ett, **A**denocarcinoma, and **D**ysplasia) to complement EAC screening and

prevention. We conclude that, pending future prospective validation, our non-invasive circulating miRNA-based signature could potentially be transformative in the clinic and improve survival outcomes.

MATERIALS AND METHODS

Study Populations

This study analyzed data from 792 patients from publicly available miRNA expression datasets (N=134) and four prospectively collected independent clinical cohorts (N=658, **Figure 1**). One clinical cohort was comprised entirely of histological bio-specimens, and three were based on blood for the development, training, and independent evaluation of the liquid biopsy assay (**Supplementary Table 1**).

The *in-silico* discovery phase utilized expression data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO, accession number: GSE164560). It was designed to identify biomarkers differentially expressed between EAC tissue and normal esophageal mucosa (N=134) and then select those demonstrating a statistically significant increase across disease stages (Anova $p < 0.05$) from patient-matched normal mucosa to LGD, HGD, and EAC (N=32). We further excluded miRNAs lacking significant expression differences between EAC tissues (Stage I-III) and patient-matched normal mucosa (N=42 each) from a separate clinical cohort (Radboud University Medical Center, Netherlands). The remaining miRNAs were examined in our *"Development cohort"* (N=108), encompassing serum specimens from 51 patients with EAC/HGD, 27 with BE/LGD, and 30 NDCs (Norton Thoracic Institute at St. Joseph's Hospital and Medical Center, Phoenix, AZ, USA; and Baylor University Medical Center, Dallas, TX, USA).

The diagnostic assay was developed in the "*Training cohort"* (N=160), which included 96 patients with EAC/HGD (Veneto Institute of Oncology IOV-IRCCS, Padova, Italy) and 64 NDCs. Finally, the blood-based assay was externally and independently tested in the "*Testing cohort"* (N=306), which included 125 patients with EAC/HGD, 98 BE/LGD, and 74 NDCs (Queen's University Belfast, United Kingdom, and the National Cancer Registry Ireland, Ireland, for the FINBAR study; The Johns Hopkins Hospital, Baltimore, MD, USA; Translational Genomics

Research Institute, Phoenix, AZ, USA; 9 excluded after quality control). Full details for study populations are in the **Supplementary Methods**.

All individuals diagnosed with EAC, HGD, LGD, or BE were considered cases. All individuals who received a negative endoscopic evaluation of the foregut during BE screening were considered NDCs [35]. The presence of dysplasia was attested and confirmed by a second pathologist with expertise in esophageal diseases [21,36].

Study Design

This study was an international, multi-institutional, retro-prospective, multi-phase biomarker study covering EDRN phases I, II, and III (Early Detection Research Network), STARD and TRIPOD-AI compliant (both as **Supplementary**) and included both tissue- and the blood-based biospecimens. Briefly, EDRN phase I is intended to discover the biomarkers associated with the condition of interest (BE/EAC, in this case), EDRN phase II is intended to develop a diagnostic assay for the condition of interest, and EDRN phase III tests the performance of the assay in an independent and external cohort (**Supplementary Figure 1**). Three independent tissue-based cohorts were utilized to discover biomarkers with diagnostic potential for EAC and BE and prioritize the biomarkers with the highest diagnostic potential (phase I). The serum biomarker panel was finalized in the blood-based biomarker *"development cohort."* Subsequently, we used a two-level machine learning approach to train an EAC/BE risk-score formula ('*EMERALD*') on qRT-PCR data from the training cohort (N=160, phase II). Finally, the model was fully locked and then tested in an independent, non-overlapping, external testing cohort (N=306, phase III).

The study was approved by the Institutional Review Boards of each participating institution (Baylor University, Johns Hopkins Hospital, National Cancer Registry Ireland, Queen's

University Belfast, Radboud University, St. Joseph's Hospital, Translational Genomics Research Institute, and Veneto Institute of Oncology; **Supplementary Methods** and **Supplementary Table 1**), conducted in accordance with the Declaration of Helsinki, and registered and completed on ClinicalTrials.gov (NCT06381583). All participants provided written informed consent. There was no patient or public involvement in the design, conduct, reporting, interpretation, or dissemination of the study.

Assay

Tissue samples were collected therapy-naïve, placed in RNAlater immediately after surgery, and stored at -80°C. Whole blood samples were collected primarily before treatment, centrifuged at 3000 g for ten minutes within 12 hours after collection, and stored in RNase-free Eppendorf tubes at -80°C.

RNA was isolated from tissue and serum using the RNeasy Mini and miRNeasy Serum/Plasma Kits, respectively (Qiagen, Valencia, CA). RNA was then reverse-transcribed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems). Real-time PCRs were conducted using MicroRNA Assay Kits and TaqMan Universal Master Mix II using QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). The expression of miRNAs was normalized against U6, a commonly used endogenous control (Ambion, Austin, TX), and the data was $2^{-\Delta Ct}$ transformed. Further details are in the **Supplementary Methods**.

Statistical Approaches

Candidate biomarkers were selected in the TCGA dataset based on the following criteria: $\log_2$(Fold-Change) >1 (EAC/HGC vs NDCs), a Benjamini-Hochberg-adjusted p-value <0.00001, an

individual candidate AUROC >70%, and an average miRNA expression level higher than the median of all differentially expressed miRNAs. This phase was adequately powered (power=0.94) to detect a 2-fold change (ρ = 200%) at a false discovery rate of 5% under conservative specifications of the depth of coverage for the transcript (λ_0=30x) and coefficient of variation in expression between samples (CV=0.5).

In all qPCR experiments, expression levels were compared using two-sided Student's *t*-tests for paired comparisons and ANOVA for comparisons between multiple groups. A p-value < 0.05 was considered statistically significant. The area under the receiver operating characteristic curves (AUROCs) with 95% confidence intervals (CI) was computed by the method of DeLong, with optimal cutoff thresholds determined by Youden's index. The odds ratios of disease as a function of the *EMERALD* score were computed with restricted cubic spline curves.

The final diagnostic model, named *EMERALD*, involved two independently trained ML models using six candidate biomarkers. The two algorithms used, XGBoost and AdaBoost, are popular models that employ a sequential iterative boosting strategy from weaker learners (decision trees for XGBoost and decision stumps for AdaBoost). Multivariate logistic regression was used to derive a formula to predict EAC risk from the two models. This model was fully locked for independent and external testing.

We employed a Markov model to simulate five cohorts of patients with chronic GERD, all aged 45 years, undergoing one of five screening options: endoscopy every 10 years, *EMERALD*-based screening every five, three, or one year, vs. no screening. The model utilizes a cycle time of one year and runs for 30 years, allowing one to observe early detection, disease prevention, and disease stage anticipation (stage shift). Quality-adjusted life years (QALYs) and costs serve as the primary outcome measures. The model incorporates the possibility of progression to BE, dysplastic BE, and, ultimately, EAC. The natural history of EAC, the compliance

with screening, the treatment outcomes, and the associated costs were derived from the literature or, when unavailable, were internally derived [37-40]. Full details on the Markov model assumptions are presented in the **Supplementary Methods** and **Supplementary Table 2**. All analyses were performed in R.

RESULTS

Characteristics of patients with EAC, LGD, and HGD, and healthy subjects

The characteristics of the study participants are summarized in **Supplementary Table 1**. There were no significant differences in the distribution of age and sex between cases and NDCs in any cohort. Tissue and serum specimens from patients with EAC, HGD, LGD, and BE, as well as healthy subjects used in this study, were primarily collected prior to surgery or chemotherapy treatment. Overall, our study included 783 individual patient-derived specimens, of whom 134 were from in-silico studies, 84 were from a tissue cohort (42 vs. 42), and 565 were from the blood cohorts (258 EAC, 139 BE, and 168 NDCs). All participants labeled as NDCs underwent upper endoscopic examination to exclude foregut diseases.

Identification of candidate miRNAs

In the first part of the tissue-based phase, we interrogated the TCGA miRNA expression dataset to identify candidate biomarkers that can distinguish patients with EAC from healthy subjects. We initially identified 22 differentially expressed miRNAs based on differential gene expression and significance level (**Figure 2A**). We then ranked them based on AUROC values and selected only the candidates that demonstrated a discriminative AUROC value of > 70% and, finally, excluded the miRNAs with a low expression level. After this initial selection, we identified a pool of 14 candidate biomarkers of potential diagnostic interest (**Figure 2B**). Applying the unweighted pair-group Ward-D2 method for unsupervised clustering, only the cancer status was co-segregated with unsupervised clustering, while other clinical characteristics (biological sex, stage, histological differentiation, and race) did not (**Figure 2C**). Next, we tested whether these biomarkers were also differentially expressed during the malignant progression from BE to LGD, HGD, and EAC. Using a public dataset (GSE16456) of 32 tissue biospecimens (6 EAC, 5 HGD, and 5 LGD with corresponding patient-matched normal mucosa), we assessed the miRNA expression

levels of these 14 candidates, quantified with microarrays (as opposed to sequencing). We observed a statistically significant trend of progressively increasing expression from normal mucosa towards EAC (Anova, $p<0.05$) for 10 of the 14 biomarkers. Because the overall aim of this study was to develop a biomarker signature capable of detecting EAC and its precursor lesions, four miRNAs were excluded (hsa-miR-135b-5p, hsa-miR-196a-1-5p, hsa-miR-335-3p, and hsa-miR-15b-5p, **Supplementary Figure 2**)

Transitioning from *in-silico* analyses to RT-qPCR, we evaluated the expression levels of the remaining 10 candidates in a clinical cohort of 42 EAC patients with patient-matched adjacent normal tissue specimens. We confirmed the robustness of the *in-silico* predictions by verifying that nine were significantly overexpressed in EAC tissues also in the first clinical cohort of our study (p-value < 0.05, two-sided paired Student's *t*-tests; **Supplementary Figure 3**). Therefore, after the exclusion of one miRNA (hsa-miR-17-5p), nine biomarker candidates were given full consideration for their ability to differentiate EAC from NDCs and were therefore carried over to the blood-based phase of our study.

In summary, the tissue-based discovery phase employed a systematic approach to identify a panel of 9 candidate miRNAs independently associated with both EAC and its precursor lesions across three quantification methods (sequencing, microarray, and RT-qPCR) in three independent cohorts.

Development of the circulating miRNA panel

Upon transitioning our tissue-based discovery phase into a blood-based assay, we sought to confirm whether the 9 tissue-derived candidate miRNAs could be measured in blood and if these were also upregulated in the serum collected from EAC patients. In our development cohort (N=108; 78 cases vs. 30 NDCs), three miRNAs (hsa-miR-181a-5p, hsa-miR-181b-5p, and hsa-miR-

196b-5p) had an expression level below the detection limit (average cycle threshold >35) and were excluded. The remaining six miRNAs (hsa-miR-106b-5p, hsa-miR-146a-5p, hsa-miR-15a-5p, hsa-miR-18a-5p, hsa-miR-21-5p, and hsa-miR-93-5p) were abundant in blood and could be carried over to the subsequent phases of our study (model training and independent testing). Interestingly, these biomarkers all demonstrated the potential to discriminate cases vs. NDCs, with AUROC values ranging from 62.1% to 80.8% (**Figure 3**). More importantly, we observed that five of these biomarkers were significantly upregulated in the EAC/HGD vs. NDC pair-wise comparison and in patients with BE/LGD compared to NDCs, supporting their potential as non-invasive biomarkers of EAC pre-cancerous lesions (p-value <0.05). Four of these biomarkers were expressed at similar levels in BE/LGD and EAC/HGD, potentially indicating that they represent neoplastic processes that occur early during esophageal malignant transformation and may be of diagnostic potential for both EAC and its precursor lesions. To assess the cancer-specificity of our six selected miRNAs, we compared their blood expression levels in two large and independent multi-cancer datasets (GSE113486, N=580, of whom N=100 NDCs and N=480 individuals with one of twelve cancers; and GSE113740, N=1334, of whom N=1033 NDCs and N=301 with one of twelve cancers). Differential gene expression analysis revealed virtually no statistically significant differences in miRNA levels in blood between individuals with various cancers and their respective controls (**Supplementary Figure 4**). This corroborates the high EAC/BE specificity for the miRNAs that were selected during the discovery phase of this study.

Development and independent testing of a circulating miRNA signature

Next, we assessed the expression of these six miRNAs in the training cohort (N=160, 96 EAC patients, of whom 23 had stage 0 disease, pT_{is}; 64 NDCs). In this cohort, too, all six miRNAs were significantly upregulated in the serum of patients compared with NDCs (p-value <0.05, two-sided Student's *t*-tests; **Supplementary Figure 5**). We then trained a stacked machine learning model

(XGBoost + AdaBoost, **Supplementary Figure 6A**) based on ΔCt values. The resulting classifier, *EMERALD*, was based on the contributions from both XGBoost and AdaBoost. The most important contributors to both models were hsa-miR-106b-5p, hsa-miR-93-5p, and hsa-miR-18a-5p (**Supplementary Figure 6B** and **6C** for XGBoost; **Supplementary Figure 6D** and **6E** for AdaBoost). This approach allowed the stratification of patients in the training cohort into high and low-risk groups based on Youden's index (positivity threshold= 3.403). As a result, the blood-based test achieved a high performance in distinguishing EAC patients from NDCs with an AUROC of 97.6% (95% CI, 95.5%-99.6%), with a corresponding sensitivity of 95.8% and a specificity of 95.2% (**Figure 4A**). Importantly, the assay's sensitivity for Tis was 95.7% ($CI_{95\%}$: 79.0 – 99.2%), 100.0% for stage I ($CI_{95\%}$: 100 – 100%), and 92.0% for Stage II EAC ($CI_{95\%}$: 75.0 – 97.8%).

To independently test the diagnostic accuracy and performance of our *EMERALD* liquid biopsy, we evaluated the signature's robustness in an external testing cohort. Among the 297 serum samples included in the testing cohort (118 EAC, 105 BE, 74 NDCs), we observed substantial replicability of our training efforts, where the 6-miRNA signature maintained a strong ability to distinguish patients from NDCs, with an AUROC value of 91.9% (**Figure 4B**) and corresponding sensitivity and specificity values of 82.5% and 90.5%, respectively (**Table 1**), including sensitivity of 85.6% for EAC ($CI_{95\%}$=78.1-90.8%) and 85.2% for non-dysplastic BE ($CI_{95\%}$=76.4–91.2%). More interestingly, the key hypothesis of our machine-learning approach was that combining two algorithms would enhance assay robustness. The density plots (**Figure 4C** and **4D**, respectively) confirm this, as the distribution of cases and NDCs in both training and testing cohorts closely resemble each other. Two distinct clusters emerge in both cohorts – one enriched with cases (green) in the top-right corner and another enriched with NDCs (pink) in the bottom-left corner.

Table 1. Statistical evaluations of the EMERALD test for differentiating cases from controls in the training and testing cohorts.

	Training cohort [N=160]			**External and independent testing cohort** [N=297]		
Area under the receiver operating characteristic curve *(95% C.I.)*		**97.6%** (95.5 – 99.6%)			**91.9%** (88.3 – 95.5%)	
	No.	**No. detected**	**Sensitivity** *(95% C.I.)*	**No.**	**No. detected**	**Sensitivity** *(95% C.I.)*
Esophageal adenocarcinoma and Barrett's esophagus, with or without dysplasia	96	92	95.8% (89.8 – 98.4%)	223	184	82.5% (77.0 – 86.9%
EAC/HGD	96	92	95.8% (89.8 – 98.4%)	125	105	84.0% (76.6 – 90.4%)
BE/LGD [A]	--	--	--	98	79	80.6% (71.7 – 87.2%)
	No.	**No. negative**	**Specificity** *(95% C.I.)*	**No.**	**No. negative**	**Specificity** *(95% C.I.)*
Non-disease controls, all	64	61	95.2% (87.1 – 98.4%)	74	67	90.5% (81.7 – 95.3%)
Non-disease controls, with long-standing or refractory heartburn [B]	--	--	--	47	44	93.6% (82.8 – 97.8%)
Non-disease controls, with dysphagia, odynophagia, or regurgitation [B]	--	--	--	39	37	94.9% (83.1 – 98.6%)
Non-disease controls, with epigastralgia [B]	--	--	--	51	50	98.0% (89.7 – 99.7%)

Footnotes: (**A**) The training cohort enrolled several Tis, but not BE/LGD; (**B**) The training cohort did not include symptom-specific data other than the indication for BE screening

It is noteworthy that while the training cohort primarily consisted of patients with early-stage or *in situ* EAC, the testing cohort also included a significant number of patients with pre-cancerous lesions (BE, LGD, or HGD). Despite this compositional difference, patients with either EAC or pre-cancerous lesions consistently displayed higher *EMERALD* values compared to NDCs in both cohorts (**Figures 5A and 5B**, respectively). Furthermore, we investigated whether higher *EMERALD* scores correlated with a greater likelihood of esophageal disease. Interestingly, in both cohorts, the odds ratios of having the disease progressively increased with higher *EMERALD* values (**Figure 5C and 5D**, respectively). The overall trend of the spline curves supports the reproducibility of the assay across both cohorts despite their differences.

Discriminatory capacity among controls with GERD-related symptoms

All NDCs met the criteria for BE screening, and their endoscopic evaluation was negative (chronic GERD or risk factors for BE and EAC [21-25]). Next, we evaluated the model's performance according to symptom clusters. In the testing cohort, 63.5% of the NDCs had long-standing, persistent, or treatment-refractory heartburn, with a few having reflux esophagitis (**Figure 6A**). The results were promising, demonstrating that *EMERALD* values were higher in cases than controls (**Figure 6B**, p<0.0001 for all comparisons, Student's t-tests). There was no statistically significant difference between GERD patients with vs. without reflux esophagitis (p=0.79) and between BE/LGD vs. EAC/HGD (p=0.45). Indeed, the circulating 6-miRNA signature maintained a robust discriminatory power for the aggregate measure of esophageal cases vs. symptomatic controls (N=270; AUROC = 95.0%, $CI_{95\%}$: 92.2 – 97.9%, sensitivity: 92.8%, specificity: 85.1%) and, most importantly, it distinguished pre-malignant lesions from symptomatic controls (n = 145; AUROC: 94.8%, $CI_{95\%}$ 91.3-98.2%, sensitivity: 92.8%, same specificity; **Figure 6C**). In addition, the performance of the 6-miRNA signature was superior to individual miRNAs when discriminating EAC/HGD from symptomatic controls (N=172, **Figure 6D**), achieving an AUROC value of 95.3% (95% CI 92.2 – 98.3%) with 93.6% sensitivity and the same specificity. Similar results were observed among those who reported symptoms other than heartburn, including epigastralgia (**Suppl. Figures 7A-7C**) or dysphagia, odinophagia, and regurgitation (**Suppl. Figures 7D-7F**). Finally, to assist clinicians in prioritizing endoscopic evaluations, we conducted a multinomial multivariate logistic regression analysis. This analysis utilized the same biomarkers to predict, among EMERALD-positive subjects, the "high-risk" (EAC/HGD) and the "low-risk" (BE/LGD) positive subjects. This analysis identified high-risk and low-risk subgroups within EMERALD-positive subjects, providing valuable information for risk stratification (**Suppl Figure 8**). Collectively, these data confirmed the diagnostic significance of our non-invasive, 6-miRNA

signature and highlighted its ability to separate symptomatic patients with pre-malignant or malignant diseases from those without.

Finally, using the sensitivity and specificity calculated in the testing cohort, we constructed a Markov-based decision model to simulate the course of events for five cohorts of patients, aged 45 at simulation start, undergoing no screening, endoscopy-based screening every 10 years, or *EMERALD*-based screening every five-, three-, or one-year for 30 years (**Figure 7A**). The results of our cost-effectiveness analysis support that both an endoscopy-first approach and a non-invasive approach would lead to a stage-shift effect, where an increase in early-stage diagnoses would be observed after a few years, with a subsequent decrease in the number of late-stage diagnoses (**Figures 7B-7E**), which would result in a reduction in mortality (**Figure 7F**). While the costs of a more compliant and non-invasive approach would initially surpass the costs of a less compliant endoscopy-first approach (especially at more frequent intervals of testing), the overall costs would be in favor of a non-invasive program after 15 years, justified by a reduction in the number of patients with more advanced-stage and more expensive diagnoses (**Figure 7G**), with a corresponding increase in QALYs (**Figures 7H**). Finally, we estimated the cost-effective frontier to be in favor of a non-invasive, *EMERALD*-based approach at 5-year intervals (**Figure 7I**), which would reduce the number of cancer-attributable deaths with a stage-shift effect for a lower price than other intervals of testing (**Supplementary Table 3**).

DISCUSSION

This multicenter study is the first to successfully develop and externally validate a liquid biopsy for the noninvasive diagnosis of EAC and BE (with and without dysplasia). We employed a multi-step approach with multiple patient cohorts to ensure our final assay could capture the full spectrum of the disease, ranging from GERD to EAC. Leveraging ML and rigorously selected biomarkers, the *EMERALD* assay was developed and independently tested in two large cohorts. The assay effectively distinguished BE with and without dysplasia (LGD/HGD), as well as EAC, from control groups, including non-disease controls and chronic GERD patients with and without esophagitis. This offers a valuable addition to the existing screening options for this aggressive and deadly cancer. Furthermore, using a Markov model, we estimated that implementing the *EMERALD* assay at a 5-year interval could offer the most cost-effective strategy to reduce EAC mortality and incidence.

Once a rare cancer, EAC has become the dominant form of esophageal cancer in developed countries [13,14]. While treatment advancements have modestly improved survival rates, half of patients die within a year of their diagnosis [15]. A key limitation of the current approach lies in its reactive nature, focusing only on EAC. A more strategic approach would prioritize identifying individuals at high risk, those with BE [16-20]. Though largely benign, BE represents a precursor lesion for EAC. Early detection and intervention at this pre-malignant stage, with dysplasia surveillance and treatment, can interrupt the disease's natural history and prevent it [26-28]. The current screening strategies could benefit from additional, non-invasive options. The Cytosponge™ represents such an option, detecting the trefoil factor 3 (TFF3) through a patient-swallowed sponge that captures esophageal epithelial cells [41]. In a pragmatic, multi-centric trial, it demonstrated encouraging patient interest rates (39% expressing interest), with high sensitivity for BE (80 to 90%, depending on the BE segment length) and specificity (92%) [42,43]. A blood-

based test expands the arsenal of screening tools by offering a readily repeatable, minimally invasive approach that might improve overall screening participation.

MiRNAs have long been recognized as promising non-invasive biomarkers due to their structural stability and abundance in circulation[44]. Studies examined the diagnostic potential of serum miRNAs in various cancers[45]. However, few attempted evaluation of the circulating miRNAs in EAC and BE, and most included analysis of single miRNAs or biomarkers not discovered in a comprehensive manner[46]. Not surprisingly, the diagnostic potential of these individual miRNA markers was limited, and the miRNA panels were constructed with biased criteria or lacked adequately powered clinical validation cohorts. Nonetheless, these studies highlighted the potential of circulating miRNAs and set the stage for more comprehensive studies [33,34]. Interestingly, three of the biomarkers included in the EMERALD assay (miR-15a-5p, miR21-5p, and miR-93-5p) were previously identified as potential biomarkers for the non-invasive diagnosis of BE/EAC by others, too [47-49]. In this multicenter study, we first analyzed independent cohorts of EAC tissues to identify the most consistently overexpressed miRNAs in EAC and HGD patients, hence ensuring their clinical relevance. Subsequently, we developed and validated a circulating miRNA signature by applying rigorous bioinformatic and statistical algorithms in multiple cohorts to validate the circulating biomarker panel. To the best of our knowledge, this is the largest cohort of patients with EAC, HGD, LGD, BE, and healthy subjects, to date, to establish and validate such an approach. Nonetheless, we would like to acknowledge that our study was retrospective in nature, and future prospective studies will allow a more rigorous assessment of this blood-based test. Staging information was not available for patients with EAC from the FINBAR study (from which most of the testing samples were derived). Therefore, while we cannot absolutely confirm the stage-specific sensitivity derived from the training cohort, we do not expect a large gap in sensitivity values between early- and late-stage EAC, given the stability of all other sensitivity values for all other elements of the disease spectrum, including BE, LGD, and HGD. The nature of our BE patient cohorts, primarily consisting of

patients with long-segment BE, also restricts the analyses to this patient population. Therefore, future studies should investigate the correlation between EMERALD scores and BE lengths (including short-segment BE), and in those with gastric intestinal metaplasia. Additionally, geographic differences within the testing cohort could potentially introduce some bias, although the large sample size of the study minimizes the impact of such factors. Finally, the EMERALD blood-based test is intended for diagnostic purposes and presently may not distinguish between BE cases who are poised to progress to EAC and those who are not, like the Cytosponge™. However, other tools are available to predict the risk of progression and may be used with EMERALD once the blood-based test returns positive results [50].

In conclusion, we have developed a non-invasive blood assay for the early detection of EAC and its precursor lesions using a comprehensive biomarker discovery approach and successfully validated its robustness using independent multi-center cohorts. Our *EMERALD* assay has the potential to transform liquid biopsy-based cancer and high-risk precancer screening of EAC patients in the future.

REFERENCES

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin* 2024;74(1):12-49. doi: 10.3322/caac.21820
2. Kroep S, Lansdorp-Vogelaar I, Rubenstein JH, et al. An Accurate Cancer Incidence in Barrett's Esophagus: A Best Estimate Using Published Data and Modeling. *Gastroenterology* 2015;149(3):577-85 e4; quiz e14-5. doi: 10.1053/j.gastro.2015.04.045
3. Lagergren J, Bergstrom R, Lindgren A, et al. Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *N Engl J Med* 1999;340(11):825-31. doi: 10.1056/NEJM199903183401101
4. Singh S, Manickam P, Amin AV, et al. Incidence of esophageal adenocarcinoma in Barrett's esophagus with low-grade dysplasia: a systematic review and meta-analysis. *Gastrointest Endosc* 2014;79(6):897-909 e4; quiz 83 e1, 83 e3. doi: 10.1016/j.gie.2014.01.009
5. Klaver E, Bureo Gonzalez A, Mostafavi N, et al. Barrett's esophagus surveillance in a prospective Dutch multi-center community-based cohort of 985 patients demonstrates low risk of neoplastic progression. *United European Gastroenterol J* 2021;9(8):929-37. doi: 10.1002/ueg2.12114
6. Kastelein F, van Olphen SH, Steyerberg EW, et al. Impact of surveillance for Barrett's oesophagus on tumour stage and survival of patients with neoplastic progression. *Gut* 2016;65(4):548-54. doi: 10.1136/gutjnl-2014-308802
7. 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(8):1252-60. doi: 10.1136/gutjnl-2014-308865
8. Visrodia K, Singh S, Krishnamoorthi R, et al. Systematic review with meta-analysis: prevalent vs. incident oesophageal adenocarcinoma and high-grade dysplasia in Barrett's oesophagus. *Aliment Pharmacol Ther* 2016;44(8):775-84. doi: 10.1111/apt.13783
9. Shawihdi M, Thompson E, Kapoor N, et al. Variation in gastroscopy rate in English general practice and outcome for oesophagogastric cancer: retrospective analysis of Hospital Episode Statistics. *Gut* 2014;63(2):250-61. doi: 10.1136/gutjnl-2012-304202
10. Dulai GS, Guha S, Kahn KL, et al. Preoperative prevalence of Barrett's esophagus in esophageal adenocarcinoma: a systematic review. *Gastroenterology* 2002;122(1):26-33. doi: 10.1053/gast.2002.30297
11. Bhat SK, McManus DT, Coleman HG, et al. Oesophageal adenocarcinoma and prior diagnosis of Barrett's oesophagus: a population-based study. *Gut* 2015;64(1):20-5. doi: 10.1136/gutjnl-2013-305506
12. Hur C, Miller M, Kong CY, et al. Trends in esophageal adenocarcinoma incidence and mortality. *Cancer* 2013;119(6):1149-58. doi: 10.1002/cncr.27834
13. Arnold M, Rutherford MJ, Bardot A, et al. Progress in cancer survival, mortality, and incidence in seven high-income countries 1995-2014 (ICBP SURVMARK-2): a population-based study. *Lancet Oncol* 2019;20(11):1493-505. doi: 10.1016/S1470-2045(19)30456-5
14. Rodriguez GM, DePuy D, Aljehani M, et al. Trends in Epidemiology of Esophageal Cancer in the US, 1975-2018. *JAMA Netw Open* 2023;6(8):e2329497. doi: 10.1001/jamanetworkopen.2023.29497
15. Liu KS, Raza SA, El-Serag HB, et al. Trends in Esophageal Adenocarcinoma and Esophageal Squamous Cell Carcinoma Incidence in the United States from 1992 to 2019. *Cancers (Basel)* 2022;14(24) doi: 10.3390/cancers14246049
16. Sami SS, Moriarty JP, Rosedahl JK, et al. Comparative Cost Effectiveness of Reflux-Based and Reflux-Independent Strategies for Barrett's Esophagus Screening. *Am J Gastroenterol* 2021;116(8):1620-31. doi: 10.14309/ajg.0000000000001336

17. Xia R, Zeng H, Liu W, et al. Estimated Cost-effectiveness of Endoscopic Screening for Upper Gastrointestinal Tract Cancer in High-Risk Areas in China. *JAMA Netw Open* 2021;4(8):e2121403. doi: 10.1001/jamanetworkopen.2021.21403
18. Rubenstein JH, Omidvari AH, Lauren BN, et al. Endoscopic Screening Program for Control of Esophageal Adenocarcinoma in Varied Populations: A Comparative Cost-Effectiveness Analysis. *Gastroenterology* 2022;163(1):163-73. doi: 10.1053/j.gastro.2022.03.037
19. Saha B, Vantanasiri K, Mohan BP, et al. Prevalence of Barrett's Esophagus and Esophageal Adenocarcinoma With and Without Gastroesophageal Reflux: A Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol* 2023 doi: 10.1016/j.cgh.2023.10.006
20. Vaughan TL, Fitzgerald RC. Precision prevention of oesophageal adenocarcinoma. *Nat Rev Gastroenterol Hepatol* 2015;12(4):243-8. doi: 10.1038/nrgastro.2015.24
21. Asge Standards Of Practice C, Qumseya B, Sultan S, et al. ASGE guideline on screening and surveillance of Barrett's esophagus. *Gastrointest Endosc* 2019;90(3):335-59 e2. doi: 10.1016/j.gie.2019.05.012
22. Fitzgerald RC, di Pietro M, Ragunath K, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* 2014;63(1):7-42. doi: 10.1136/gutjnl-2013-305372
23. Shaheen NJ, Falk GW, Iyer PG, et al. ACG Clinical Guideline: Diagnosis and Management of Barrett's Esophagus. *Am J Gastroenterol* 2016;111(1):30-50; quiz 51. doi: 10.1038/ajg.2015.322
24. Shaheen NJ, Weinberg DS, Denberg TD, et al. Upper endoscopy for gastroesophageal reflux disease: best practice advice from the clinical guidelines committee of the American College of Physicians. *Ann Intern Med* 2012;157(11):808-16. doi: 10.7326/0003-4819-157-11-201212040-00008
25. NICE. 2019 surveillance of gastro-oesophageal reflux disease and dyspepsia in adults: investigation and management (NICE guideline CG184). London2019.
26. Shaheen NJ, Sharma P, Overholt BF, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. *N Engl J Med* 2009;360(22):2277-88. doi: 10.1056/NEJMoa0808145
27. Sharma P, Shaheen 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(3):760-69. doi: 10.1053/j.gastro.2019.09.051
28. 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(12):1209-17. doi: 10.1001/jama.2014.2511
29. Sijben J, Peters Y, van der Velden K, et al. Public acceptance and uptake of oesophageal adenocarcinoma screening strategies: A mixed-methods systematic review. *EClinicalMedicine* 2022;46:101367. doi: 10.1016/j.eclinm.2022.101367
30. Wichmann RM, Fernandes FT, Chiavegatto Filho ADP, et al. Improving the performance of machine learning algorithms for health outcomes predictions in multicentric cohorts. *Sci Rep* 2023;13(1):1022. doi: 10.1038/s41598-022-26467-6
31. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer* 2015;15(6):321-33. doi: 10.1038/nrc3932
32. Miyoshi J, Zhu Z, Luo A, et al. A microRNA-based liquid biopsy signature for the early detection of esophageal squamous cell carcinoma: a retrospective, prospective and multicenter study. *Mol Cancer* 2022;21(1):44. doi: 10.1186/s12943-022-01507-x
33. Craig MP, Rajakaruna S, Paliy O, et al. Differential MicroRNA Signatures in the Pathogenesis of Barrett's Esophagus. *Clin Transl Gastroenterol* 2020;11(1):e00125. doi: 10.14309/ctg.0000000000000125
34. Petrick JL, Pfeiffer RM, Liao LM, et al. Circulating MicroRNAs in Relation to Esophageal Adenocarcinoma Diagnosis and Survival. *Dig Dis Sci* 2021;66(11):3831-41. doi: 10.1007/s10620-020-06740-2

35. Anaparthy R, Gaddam S, Kanakadandi V, et al. Association between length of Barrett's esophagus and risk of high-grade dysplasia or adenocarcinoma in patients without dysplasia. *Clin Gastroenterol Hepatol* 2013;11(11):1430-6. doi: 10.1016/j.cgh.2013.05.007
36. Weusten B, Bisschops R, Dinis-Ribeiro M, et al. Diagnosis and management of Barrett esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2023;55(12):1124-46. doi: 10.1055/a-2176-2440
37. Desai TK, Krishnan K, Samala N, et al. The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut* 2012;61(7):970-6. doi: 10.1136/gutjnl-2011-300730
38. Tramontano AC, Chen Y, Watson TR, et al. Esophageal cancer treatment costs by phase of care and treatment modality, 2000-2013. *Cancer Med* 2019;8(11):5158-72. doi: 10.1002/cam4.2451
39. Thein HH, Jembere N, Thavorn K, et al. Estimates and predictors of health care costs of esophageal adenocarcinoma: a population-based cohort study. *BMC Cancer* 2018;18(1):694. doi: 10.1186/s12885-018-4620-2
40. Menezes A, Tierney A, Yang YX, et al. Adherence to the 2011 American Gastroenterological Association medical position statement for the diagnosis and management of Barrett's esophagus. *Dis Esophagus* 2015;28(6):538-46. doi: 10.1111/dote.12228
41. Kadri SR, Lao-Sirieix P, O'Donovan M, et al. Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: cohort study. *BMJ* 2010;341:c4372. doi: 10.1136/bmj.c4372
42. Fitzgerald RC, di Pietro M, O'Donovan M, et al. Cytosponge-trefoil factor 3 versus usual care to identify Barrett's oesophagus in a primary care setting: a multicentre, pragmatic, randomised controlled trial. *Lancet* 2020;396(10247):333-44. doi: 10.1016/S0140-6736(20)31099-0
43. 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(1):e1001780. doi: 10.1371/journal.pmed.1001780
44. Redova M, Sana J, Slaby O. Circulating miRNAs as new blood-based biomarkers for solid cancers. *Future Oncol* 2013;9(3):387-402. doi: 10.2217/fon.12.192
45. Komatsu S, Ichikawa D, Hirajima S, et al. Plasma microRNA profiles: identification of miR-25 as a novel diagnostic and monitoring biomarker in oesophageal squamous cell carcinoma. *Br J Cancer* 2014;111(8):1614-24. doi: 10.1038/bjc.2014.451
46. Zhang K, Wu X, Wang J, et al. Circulating miRNA profile in esophageal adenocarcinoma. *Am J Cancer Res* 2016;6(11):2713-21.
47. Bus P, Kestens C, Ten Kate FJ, et al. Profiling of circulating microRNAs in patients with Barrett's esophagus and esophageal adenocarcinoma. *J Gastroenterol*. 2016;51(6):560-70. doi: 10.1007/s00535-015-1133-5
48. Xie Z, Chen G, Zhang X, et al. Salivary microRNAs as promising biomarkers for detection of esophageal cancer. *PLoS One*. 2013;8(4):e57502. doi: 10.1371/journal.pone.0057502
49. Bansal A, Hong X, Lee IH, et al. Serum Exosomal MicroRNA Expression Can Be a Novel Non-Invasive Strategy for the Screening of Barrett's Esophagus. *Gastroenterology*. 2013; 144(5):S-684. doi: 10.1016/S0016-5085(13)62541-X
50. Parasa S, Vennalaganti 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(5):1282-1289. doi: 10.1053/j.gastro.2017.12.009IF

FIGURE LEGENDS

Figure 1: Consort diagram for study cohorts allocation.

Figure 2: *In silico* discovery and prioritization of candidate miRNAs (**A**) Volcano plot of differentially expressed microRNAs between cases and controls using TCGA miRNA expression dataset. Color grading follows significance. 22 miRNAs were differentially expressed and 14 miRNAs (identified as fully colored) were prioritized as our initial candidates for further analysis. (**B**) Ridgeline plot of the initial pool of 14 microRNAs; (**C**) A heatmap with unsupervised clustering illustrates the expression levels of the 14 candidate miRNAs in the TCGA miRNA expression dataset.

Figure 3: Assessment of the 6 miRNAs diagnostic potential in the US (*development*) cohort. Six circulating miRNAs were significantly upregulated in cases (EAC, HGD, LGD, and BE) compared to NDCs and demonstrated AUROC values of diagnostic interest, ranging from 62.1% to 80.8%. In sub-group analysis (inserts), all candidate miRNAs demonstrated a significant differential expression between NDCs (blue) vs. EAC/HGD (maroon), five for NDCs vs. BE/LGD (orange), and two for BE/LGD vs. EAC/HGD (* $p<0.05$, ** $p<0.01$, *** $p<0.001$; two-sided Student's t-tests).

Abbreviations: AUROC, area under the receiver operating characteristic curve; Spe, Specificity; Sen, sensitivity

Figure 4: Establishment and external testing of the 6-circulating-miRNA signature (**A**) AUROC and performance metrics of the two first-level classifiers (XGB: XGBoost, orange; ADA, AdaBoost, black) and the resulting stacked model (*EMERALD*) in the Italian *(training*) cohort. (**B**) AUROC and performance metrics of the two first-level classifiers (XGB: XGBoost, orange; ADA, AdaBoost, black) and the resulting stacked model (*EMERALD*) in the UK/Irish *(testing*) cohort; (**C, D**) Density-scatter plots: the background color gradient is a density plot and highlights the areas where most observations are encountered, without knowledge of them being cases and controls. The dots (green for cases and pink for NDCs) represent the scatter plot. The same patterns can be seen in both the Italian (*training,* panel **C**) and UK/Irish (*testing,* panel **D**) cohorts, with two clusters where cases and controls appear to separate (top right and bottom left, respectively).

Figure 5: Evaluation of the diagnostic performance of the 6-circulating-miRNA signature. (**A,B**) Raincloud plots with super-imposed box and whisker plots demonstrating the distribution of *EMERALD* values between cases and controls in the Italian (*training,* panel **A**) and UK/Irish (*testing,* panel **B**) cohorts; (**C,D**) Odds ratios for the presence of a compound endpoint of EAC, HGD, LGD, or BE with restricted cubic splines in the Italian (*training,* panel **C**) and UK/Irish (*testing,* panel **D**) cohorts.

Abbreviations: EC, Esophageal cases; NDC, Non-disease controls

Figure 6: Evaluation of the *EMERALD* assay among symptomatic controls with GERD. (**A**) Waterfall plot of the *EMERALD* values in the UK/Irish (*testing*) cohort, with controls sub-labeled by the presence of symptoms; (**B**) Violin plot of *EMERALD* values according to the presence of GERD-related symptoms, with and without reflux esophagitis vs. cases with BE or LGE and cases with EAC or HGD (**** p<0.0001; two-sided Student's t-tests). (**C,D**) AUROCs of the two first-level classifiers (XGB: thin line; ADA, dashed line) and the stacked model (thick line with 95% confidence intervals) in a sub-analysis of the UK/Irish *(testing*) cohort which only included symptomatic controls vs. Barrett's esophagus cases, with or without low-grade dysplasia (**C**) and cases with either high-grade dysplasia or adenocarcinoma (**D**).

Abbreviations: BE, Barrett's esophagus; GERD, Gastroesophageal reflux disease: RE, Reflux esophagitis; LGD, Low-grade Dysplasia; HGD, High-grade dysplasia; EAC, Esophageal Adenocarcinoma; AUROC, Area under the receiver operating characteristic curve

Figure 7: Markov simulation. (**A**) Markov model structure: for each strategy, the Markov chain assumes a progression through pre-malignant disease, malignant disease, cancer recovery, recurrence, and, eventually death. Mortality can be from any cause (cardiovascular, other cancers, etc.) or stage-specific mortality rates. Early detection of Barrett's esophagus with dysplasia is assumed to trigger its eradication. EAC diagnosis, whether due to symptoms, endoscopic screening uptake, or non-invasive screening uptake, is assumed to trigger treatment. Costs and utility values are derived from the literature or, if unavailable, from internal estimates. The starting condition of each individual in the Markov chain is a 45-year-old with chronic GERD who is followed up for 30 years or until death. Five screening strategies are tested: non-invasive, *EMERALD*-based screening at 45% compliance every 5, 3, or 1 year(s), endoscopy-based screening at 10% compliance every 10-15 years, or no screening. (**B**) Number of stage I EAC diagnoses per screening strategy. (**C**) Number of stage II EAC diagnoses per screening strategy. (**D**) Number of stage III EAC diagnoses per screening strategy. (**E**) Number of stage IV EAC diagnoses per screening strategy. (**F**) Number of deaths from all causes (both EAC and non-EAC related) per screening strategy. (**G**) Annual costs associated with each screening strategy. (**H**) Quality-adjusted life years per year by screening strategy, with a call-out box to demonstrate the effects of different screening strategies. (**I**) Incremental costs and effectiveness of each screening strategy against the most cost-effective screening strategy and interval.

Abbreviations: EAC, Esophageal adenocarcinoma; USD, U.S. Dollars

A LIQUID BIOPSY TO IDENTIFY BARRETT'S ESOPHAGUS, DYSPLASIA, AND ESOPHAGEAL ADENOCARCINOMA: THE *EMERALD* MULTI-CENTER STUDY.

Jinsei Miyoshi[1,2,3] *, Alessandro Mannucci[4,5] *, Marco Scarpa[6], Feng Gao[7], Shusuke Toden[1], Timothy Whitsett[8], Landon J. Inge[9], Ross M. Bremner[9], Tetsuji Takayama[2], Yulan Cheng[10], Teodoro Bottiglieri[11], Iris D Nategaal[12], Martha J. Shrubsole[13], Ali H. Zaidi[14], Xin Wang[15,16,17], Helen G Coleman[18], Lesley A. Anderson[19], Stephen J. Meltzer[10], and Ajay Goel[1,4,20], on behalf of the FINBAR-EMERALD collaborative group

1. Center for Gastrointestinal Research; Center from Translational Genomics and Oncology, Baylor Scott & White Research Institute and Charles A. Sammons Cancer Center, Baylor University Medical Center, Dallas, TX, USA
2. Department of Gastroenterology and Oncology, Tokushima University Graduate School of Biomedical Sciences, Tokushima, Japan
3. Department of Gastroenterology, Kawashima Hospital, Tokushima, Japan.
4. Department of Molecular Diagnostics and Experimental Therapeutics, Beckman Research Institute of City of Hope, Monrovia, California, USA
5. Gastroenterology and Gastrointestinal Endoscopy Unit, Vita-Salute San Raffaele University, IRCCS San Raffaele Hospital, Vita-Salute San Raffaele University, Milan, Italy
6. Department of Surgical, Oncological and Gastroenterological Sciences, University of Padova, Padova, Italy
7. The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
8. Cancer and Cell Biology Division, The Translational Genomics Research Institute (TGen), Phoenix, AZ, USA
9. Norton Thoracic Institute, St Joseph's Hospital and Medical Center, Phoenix, Arizona
10. Division of Gastroenterology and Hepatology, Department Of Medicine And Sidney Kimmel Comprehensive Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, MD, USA
11. Baylor Scott & White Research Institute, Institute of Metabolic Diseases, Dallas, TX, USA
12. Radboud University Medical Centre, Department of Pathology, Nijmegen, the Netherlands
13. Department of Medicine, Division of Epidemiology, Vanderbilt University School of Medicine, Nashville, TN, USA
14. Allegheny Health Network Cancer Institute, Allegheny Health Network, Pittsburgh, PA, USA
15. Department of Surgery, The Chinese University of Hong Kong, Hong Kong, China
16. Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, Hong Kong, China
17. Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China
18. Cancer Epidemiology Research Group, Centre for Public Health, Queen's University Belfast, Belfast, United Kingdom
19. Centre for Health Data Science, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK
20. City of Hope Comprehensive Cancer Center, Duarte, California, USA

***** *Denotes shared first authorship*

SUPPLEMENTARY METHODS

Supplementary Study populations

This study included 792 patients, both from publicly available miRNA expression datasets (N=134) and four independent clinical cohorts (N=658). One cohort was comprised entirely of histological bio-specimens, and three were based on blood for the development, training, and independent evaluation of a liquid biopsy assay (**Supplementary Table 1 and Figure 1**).

Supplementary *In-silico* discovery cohorts

Initially, we interrogated The Cancer Genome Atlas (TCGA) miRNA expression dataset for biomarker discovery. The TCGA dataset consisted of tissue-based small RNA-seq data from 89 stage I-IV EAC patients and 13 normal esophageal mucosae. The level-3 expression data of 2082 miRNAs and the corresponding clinical data were downloaded from The Firehose Broad GDAC portal (http://gdac.broadinstitute.org/). The miRNA expression levels, measured by reads per million miRNAs mapped (RPM), were $\log_2$-transformed and then used for subsequent analyses. Given the sample size, this phase of our study was adequately powered (power=0.94) to detect a 2-fold change (ρ = 200%) at a false discovery rate of 5% under conservative specifications of the depth of coverage for the transcript (λ_0=30x) and coefficient of variation in expression between samples (CV=0.5). Another public tissue miRNA dataset was downloaded from Gene Expression Omnibus (GEO, accession number: GSE16456), with 32 paired diseased and normal tissues from 16 patients diagnosed with EAC, BE with HGD, or BE with LGD to evaluate which candidate microRNAs increase from normal, to LGD, HGD, and EAC and, hence, tracked with the neoplastic transition events [1].

Supplementary Clinical Cohorts Characteristics

To evaluate the potential of these candidate miRNAs as specific biomarkers for EAC and its precursor lesions, we assessed their expression levels and confirmed their *bona fide* biomarker status in four patient cohorts encompassing 649 clinical specimens. All cohorts remained separate during the entire analytical process to preserve their independence, and no data partitioning was involved. This evaluation involved three steps: first, we evaluated the expression of the biomarkers discovered during the *in-silico* discovery phase in a tissue cohort using RT-qPCR. Second, in a blood-based cohort, we identified which tissue markers were also present in blood. Finally, two additional blood cohorts were used to develop and test a diagnostic assay for EAC and its precursor lesions.

For tissue validation of the candidate miRNAs identified in the TCGA and GSE16456 datasets, we used a single-center case-only study conducted in the Netherlands in patients who had received

esophagectomy for EAC (Ethical Committee approval obtained from the Radboud University Medical Centre). This cohort included 42 EAC tissue slides, and 42 matched normal mucosal tissues collected from stage I-III EAC patients undergoing esophageal resection without pre-operative therapy. All participants in this cohort were enrolled at the Radboud University Medical Center, Netherlands, between 1999 and 2009.

The candidate miRNAs identified from the tissue cohort were first examined in our *"Development cohort"* (N=108), which was a multi-institutional case-control study conducted at the Norton Thoracic Institute at St. Joseph's Hospital and Medical Center, Phoenix, AZ, USA, and the Baylor University Medical Center, Dallas, TX, USA (Ethical Committee approval obtained from both institutions separately). This study phase included serum specimens from 51 EAC/HGD patients, 27 from BE/LGD patients, and 30 NDCs. All study participants were enrolled between 2011 and 2015 at the Norton Thoracic Institute at St. Joseph's Hospital and Medical Center and Baylor University Medical Center in the USA.

After the final panel of serum biomarkers was established, a diagnostic assay was developed in our "*Training cohort"* (N=160). This study cohort was conducted in Italy (Ethical Committee approval obtained from the University of Padua), and it included 96 specimens from EAC/HGD patients, all enrolled between 2011 and 2015 at the Veneto Institute of Oncology IOV-IRCCS, Padova, Italy, and 64 from NDCs. Finally, the miRNA signature was validated using our "*Testing cohort"* (N=306), a multinational case-control study that drew participants from several institutions and leveraged samples from the FINBAR study. Our *"Testing cohort"* utilized 297 serum specimens from 125 EAC/HGD patients, 98 BE/LGD patients, and 74 NDC subjects from different institutes in the UK, Ireland (FINBAR), and the USA. The FINBAR (Factors Influencing the Barrett's Adenocarcinoma Relationship) Study was an all-Ireland population-based case-control study for the etiology of EE, BE, and EAC [2,3]. Incident cases of EAC, BE, and population NDCs were recruited from Northern Ireland and the Republic of Ireland between March 2002 and July 2004, while EE cases were recruited from Northern Ireland only between 2004 and 2005. BE, EE, and NDCs were matched within 5-year age and sex strata to EAC (maximum age: 85 years). This study was compliant with the Declaration of Helsinki, and all procedures were approved by the Research Ethics Committee of the Queen's University Belfast, Northern Ireland, the Clinical Research Ethics Committee of Cork Teaching Hospitals, and the Research Ethics Committee Board of St. James's Hospital, Dublin. All the subjects provided written informed consent to participate in the study. Specifically for the *EMERALD* study, 206 patient-derived samples were drawn from the FINBAR study (EAC, N=118; BE, N=88), and all were recruited from Northern Ireland and the Republic of Ireland. The remaining 91 subjects were derived from another multi-institutional case-control study that involved the Johns Hopkins Hospital,

Baltimore, MD, and the Translational Genomics Research Institute, Phoenix, AZ (Ethical Committee approval obtained from both institutions separately). The decision to use the FINBAR cohort for external testing was based on the understanding that early detection of BE, a precursor to EAC, is crucial for preventing the progression of this preventable disease. This allowed us to accurately assess EMERALD's performance for both BE and EAC.

All subjects in this study received an endoscopic examination of their foregut for BE screening. The indication for BE screening was given based on the local guidelines (ASGE/ESGE). Additionally, among the *Testing cohort* NDCs, 63.5% complained of persistent or treatment-refractory heartburn, 68.9% complained of epigastralgia, and 52.7% complained of dysphagia, odynophagia, or regurgitation, while 12.2% also reported other symptoms (sore throat, cough, Globus, nausea). Moreover, 5.4% had a family history of BE for which they received an indication for BE screening, but the examination was negative. All subjects were enrolled between 2002 and 2004. In total, only 9 samples in the *"Testing cohort"* were excluded from the analyses by quality control (QC).

The study was approved by the Institutional Review Boards of each participating institution (previously reported [4-6] and detailed in **Supplementary Table 1**), conducted in accordance with the Declaration of Helsinki, and registered and completed on ClinicalTrials.gov (NCT06381583). All participants provided written informed consent.

Supplementary Study Design

Our study was STARD and TRIPOD-AI compliant (both in **Supplementary Material**) and encompassed two major phases, one tissue- and the other blood-based, respectively (**Supplementary Figure 1**). The tissue phase consisted of two sub-studies (both belonging to the EDRN phase I of this study): an *in-silico* approach to biomarker discovery in two cohorts and a clinical validation for biomarker prioritization. The blood-based segment consisted of three independent phases in three non-overlapping patient cohorts: blood-phase biomarker panel *development* (which belongs to EDRN phase I of this study), and subsequent liquid biopsy *training* (EDRN phase II), and independent *testing* (EDRN phase III).

During the *in-silico* phase, tissue-based genome-wide miRNA expression datasets from TCGA and GSE16456 were used to identify miRNAs that were overexpressed in the EAC tissue specimens compared to normal mucosa. Candidate miRNAs were then prioritized using the following criteria: $\log_2$(Fold-Change) >1 (EAC/HGC vs non-disease controls, NDCs), a strict statistical level of significance of Benjamini-Hochberg-adjusted p-value <0.00001, an individual candidate AUC >70%, and an average miRNA expression level higher than the median of all differentially expressed miRNAs. Because our intention was

to generate an assay capable of detecting both EAC and its precursor lesions, the expression levels of the candidate miRNAs were then evaluated in a second, microarray-based tissue phase that utilized a cohort of EAC, HGD, and LGD with matching normal mucosa. To mitigate the risk of cross-platform bias and batch effects, we harmonized the GSE16456 dataset using a standardized scoring method (a commonly used and widely adopted method for adjusting differences between analytical techniques). Only the candidate biomarkers that demonstrated a statistically significant (Anova $p < 0.05$) progressive increase in expression from normal to LGD, HGD, and EAC were carried over to the following phase of tissue validation. To confirm the results from sequencing and microarray methods, we further tested the differential expression of the remaining candidates also by RT-qPCR in a cohort of matched tumors and adjacent normal tissues from 42 EAC patients.

In order to test whether miRNAs overexpressed in tissues were also detectable and differentially expressed in blood, we assessed their expression levels and potential as biomarkers in a serum cohort of 51-EAC/HGD, 27 LGD/BE, and 30 NDC. To assess the specificity of our biomarker candidates for BE/EAC, we evaluated their expression levels in two independent and large cohorts of individuals with twelve different cancers (GSE113486 and GSE113740). This final biomarker selection step concluded the EDRN phase I of this study. Next, the efficacy of the candidate miRNAs was examined in the training (EDRN phase II) and independent validation phases (EDRN phase III). To establish an EAC risk-score formula, we employed a two-level machine learning approach (details provided in subsequent sections) using qRT-PCR data from the training cohort (N=160, 96 cases vs. 64 NDCs). After the efforts to prioritize biomarkers that would detect both EAC and its precursor lesions, the choice to train an algorithm primarily on EAC/HGD was strategic to ensure that the resulting diagnostic model would be particularly effective in detecting the highest risk lesions. After training, the model was fully locked and validated in an independent, non-overlapping, external cohort (N=306), which included, after QC, 125 EAC/HGD, 98 BE/LGD, and 74 NDCs, where performance and robustness were evaluated.

Supplementary Definitions and inclusion/exclusion criteria

All individuals diagnosed with EAC, HGD, LGD, or BE were considered cases. All individuals with a negative endoscopic evaluation of the foregut were considered NDCs [7-9]. Because this study aimed to develop a test that would complement BE screening, all the NDCs were individuals whose screening for BE was recommended according to local guidelines (ASGE/ESGE) based on their risk of having BE and whose endoscopic examination ruled out BE and EAC. Individuals without BE/EAC but with non-erosive gastroesophageal reflux, reflux esophagitis, or hiatal hernia were considered NDCs. The presence or

absence of low- or high-grade dysplasia was attested and confirmed by a second pathologist with expertise in esophageal diseases [10,11].

Supplementary RNA isolation and quantitative real time-PCR

Tissue samples from our clinical tissue cohort (tumor and the corresponding normal mucosa) were obtained from patients subjected to esophagectomy without any pre-operative therapy, and they were immediately placed in RNAlater (Qiagen, Germany) and then stored at -80°C. Whole blood samples from each participant were collected primarily before treatment and centrifuged at 3000 g for ten minutes within 12 hours after collection. The serum samples were stored in an RNase-free Eppendorf tube at -80°C.

RNA was isolated from tissue specimens using the RNeasy Mini Kit (Qiagen, Valencia, CA), and for serum RNA isolation, the miRNeasy Serum/Plasma Kit (Qiagen) kit was used according to the manufacturer's instructions to extract RNA enriched in small RNAs. Briefly, serum samples were thawed on ice and centrifuged at 10,000 rpm for five minutes to remove cellular debris. 200 µL of supernatant was lysed in 1,000 µL of Qiazol Lysis Reagent. For normalization of sample-to-sample variation during the RNA isolation procedures, synthetic *C. elegans* miRNA (cel-miR-39, Qiagen) was added to each denatured sample[12]. For miRNA-based qRT-PCR assays, RNA from tissue and serum specimens was reverse-transcribed using the TaqMan MicroRNA Reverse Transcription Kit (Applied Biosystems) according to manufacturer's instructions in a total reaction volume of 6 µL. Real-time PCRs were conducted using MicroRNA Assay Kits and TaqMan Universal Master Mix II, no UNG (Applied Biosystems) using QuantStudio 6 Flex Real-Time PCR System (Applied Biosystems). The expression of miRNAs was normalized using U6 (Ambion, Austin, TX), an endogenous control commonly used to normalize miRNA expression data[13] [14]. All data were represented as $2^{-\Delta Ct}$.

Supplementary Model Architecture and Hyperparameters

The final diagnostic model, named *EMERALD*, is a stacked model. Two independent ML models were allowed to learn patterns from the final six candidate miRNAs. The two algorithms used, XGBoost and AdaBoost, are popular models that employ a sequential iterative boosting strategy to convert weaker learners (decision trees for XGBoost and decision stumps for AdaBoost) into a robust model. XGBoost is generally preferred for complex, high-dimensional data due to its extensive system optimization options and faster computations. However, AdaBoost performs well for data with low noise levels, as was the case in our dataset after our consistent efforts in candidate miRNA selection. As both approaches allow for

extracting meaningful information from data, but in different ways, both were used independently. In the second step, multivariate logistic regression was employed to derive a formula to predict EAC risk from the two models (‘glm’ function of R) and, at this time, the resulting model was fully locked before independent and external validation.

The XGBoost model was allowed to undergo a maximum of 500 training rounds with gradient boosting trees implemented by the 'xgboost' package in R (Version 0.1.3). No model updating or class imbalance methods were allowed after evaluation. Given our goal to optimize both sensitivity and specificity, we utilized the area under the curve (AUC) and precision-recall (‘aucpr’ function) as the evaluation metrics. To balance model complexity and avoid overfitting, we allowed the tree depth to reach a maximum of five branches, imposed a 75% subsample parameter for training data selection, and employed a high-pruning strategy (γ=5). Finally, we determined model explainability by analyzing feature importance and SHAP values using the 'SHAPforxgboost' (Version 0.1.3) and 'fmsb' (Version 0.7.6) packages in R. The AdaBoost model, implemented with the 'ada' package in R (Version 2.0-5), underwent up to 50 training iterations with a sampling fraction of 55% and, to control overtraining, the learning rate was reduced to 20% with the default boosting strategy. For this model, too, we analyzed feature importance and SHAP values, using the ‘varplot’ function within the 'ada' package and the ‘kernelshap’ package (Version 0.4.1) in R, respectively. XGBoost outputs were probabilities ranging from 0 to 1, AdaBoost outputs were dichotomous 1 or 2 outcomes, and GLM outputs were log odds.

Supplementary Statistical Analysis

During the *in-silico* biomarker discovery phase, differential miRNA expression analysis was performed using an empirical Bayes method by ‘limma’ package in R, and the resulting p-values were adjusted using the Benjamini-Hochberg method. For visual processing of sequencing data, volcano plots, heatmaps, and ridgeline plots were generated in R using the ‘EnhancedVolcano’ (Release 3.18), ‘pheatmap’ (Version 1.0.12), and ‘ggridges’ (Version 0.5.6) packages, respectively. The statistical power for identification of differentially expressed miRNAs from the TCGA dataset was calculated using the Bioconductor ‘RNASeqPower’ package in R (Version 2.12.0).

In all qPCR experiments, differential miRNA expression was analyzed using two-sided Student’s *t*-tests for paired comparisons and ANOVA for comparisons between multiple groups, with a p-value of < 0.05 considered statistically significant. A receiver operating characteristic (ROC) curve was generated, and the area under the ROC curve (AUC) with 95% confidence intervals (CI) was computed by the method of DeLong with 2000 stratified bootstrap replicates ('pROC,' Version 1.18.5), with optimal cutoff

thresholds determined by Youden's index ('cutpointr,' Version 1.1.2). The odds ratios of disease as a function of the *EMERALD* score were computed with restricted cubic spline curves ('plotRCS', Version 0.1.4). Finally, to assist clinicians in prioritizing endoscopic evaluations, we conducted an additional analysis to determine whether the same biomarkers used in EMERALD could differentiate high-risk (EAC/HGD) from low-risk (BE/LGD) individuals within the EMERALD-positive population. To achieve this, we performed a multivariate logistic regression analysis, generating a risk score based on these biomarkers [y = 0.548 + (-0.004***hsa-miR-106b-5p*) + (-0.041***hsa-miR-146a-5p*) + (-0.07***hsa-miR-15a-5p*) + (0.092***hsa-miR-18a-5p*) + (-0.034***hsa-miR-21-5p*) + (0.075***hsa-miR-93-5p*)].

We employed a Markov model-based decision analysis to evaluate the potential clinical and economic impact of the *EMERALD* assay. This approach simulates the course of events for five cohorts of patients with chronic GERD, all aged 45 years at the simulation start. Each cohort underwent one of five screening options: endoscopy every 10 years or *EMERALD*-based screening every five, three, or one year (implemented using the 'heemod' package, Version 1.0.1). The model utilizes a cycle time of one year and runs for 30 years, allowing sufficient time to observe the impact of early detection, disease prevention, and disease stage anticipation at presentation (stage shift). Quality-adjusted life years (QALYs) and associated costs serve as the primary outcome measures. The base-case patient is a 45-year-old with chronic reflux for at least five years. The model incorporates the possibility of progression to BE, dysplastic BE, and, ultimately, EAC. The annual incidence rates for BE (3%) and malignant transformation (0.3%), as well as the prevalence of BE at simulation start (8%), were derived from the literature [15-17]. Age-specific mortality rates for the general population were obtained from the WHO mortality registry. We opted to focus on a single age group to limit model complexity and because of limitations in high-quality data for assigning precise age-related risks. A comprehensive PubMed search identified relevant articles on the natural history of EAC, including treatment outcomes and associated costs for both conservative and surgical approaches (**Supplementary Table 3**) [9,18-21]. Data from these sources informed the model's parameters regarding progression rates, cure rates, clinical characteristics, costs, and QALYs. Internally derived costs were used when data was unavailable [9,18-21]. Importantly, the model accounts for the anticipated differences in patient compliance between screening strategies. Compliance with endoscopy was estimated at 10% based on prior studies [22-24]. For the non-invasive *EMERALD* assay, we conservatively estimated a 45% compliance rate, drawing from data on adherence observed in other liquid biopsy studies [25,26]. All analyses were performed in R.

SUPPLEMENTARY METHODS REFERENCES

1. Yang H, Gu J, Wang KK, et al. MicroRNA expression signatures in Barrett's esophagus and esophageal adenocarcinoma. *Clin Cancer Res* 2009;15(18):5744-52. doi: 10.1158/1078-0432.CCR-09-0385
2. 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(10):1585-94. doi: 10.3748/wjg.v13.i10.1585
3. Mulholland HG, Murray LJ, Anderson LA, et al. Vitamin D, calcium and dairy intake, and risk of oesophageal adenocarcinoma and its precursor conditions. *Br J Nutr* 2011;106(5):732-41. doi: 10.1017/S0007114511000742
4. Anderson LA, Murphy SJ, Johnston BT, et al. Relationship between Helicobacter pylori infection and gastric atrophy and the stages of the oesophageal inflammation, metaplasia, adenocarcinoma sequence: results from the FINBAR case-control study. *Gut* 2008;57(6):734-9. doi: 10.1136/gut.2007.132662
5. Kubo A, Cook MB, Shaheen NJ, et al. Sex-specific associations between body mass index, waist circumference and the risk of Barrett's oesophagus: a pooled analysis from the international BEACON consortium. *Gut* 2013;62(12):1684-91. doi: 10.1136/gutjnl-2012-303753
6. Anderson LA, Cantwell MM, Watson RG, et al. The association between alcohol and reflux esophagitis, Barrett's esophagus, and esophageal adenocarcinoma. *Gastroenterology* 2009;136(3):799-805. doi: 10.1053/j.gastro.2008.12.005
7. Kambhampati S, Tieu AH, Luber B, et al. Risk Factors for Progression of Barrett's Esophagus to High Grade Dysplasia and Esophageal Adenocarcinoma. *Sci Rep* 2020;10(1):4899. doi: 10.1038/s41598-020-61874-7
8. Anaparthy R, Gaddam S, Kanakadandi V, et al. Association between length of Barrett's esophagus and risk of high-grade dysplasia or adenocarcinoma in patients without dysplasia. *Clin Gastroenterol Hepatol* 2013;11(11):1430-6. doi: 10.1016/j.cgh.2013.05.007
9. Parasa S, Vennalaganti 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(5):1282-89 e2. doi: 10.1053/j.gastro.2017.12.009
10. Weusten B, Bisschops R, Dinis-Ribeiro M, et al. Diagnosis and management of Barrett esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2023;55(12):1124-46. doi: 10.1055/a-2176-2440
11. Asge Standards Of Practice C, Qumseya B, Sultan S, et al. ASGE guideline on screening and surveillance of Barrett's esophagus. *Gastrointest Endosc* 2019;90(3):335-59 e2. doi: 10.1016/j.gie.2019.05.012
12. Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A* 2008;105(30):10513-8. doi: 10.1073/pnas.0804549105
13. Shell S, Park SM, Radjabi AR, et al. Let-7 expression defines two differentiation stages of cancer. *Proc Natl Acad Sci U S A* 2007;104(27):11400-5. doi: 10.1073/pnas.0704372104
14. Corney DC, Flesken-Nikitin A, Godwin AK, et al. MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. *Cancer Res* 2007;67(18):8433-8. doi: 10.1158/0008-5472.CAN-07-1585
15. Balasubramanian G, Singh M, Gupta N, et al. Prevalence and predictors of columnar lined esophagus in gastroesophageal reflux disease (GERD) patients undergoing upper endoscopy. *Am J Gastroenterol* 2012;107(11):1655-61. doi: 10.1038/ajg.2012.299
16. Desai TK, Krishnan K, Samala N, et al. The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut* 2012;61(7):970-6. doi: 10.1136/gutjnl-2011-300730

17. Bhat SK, McManus DT, Coleman HG, et al. Oesophageal adenocarcinoma and prior diagnosis of Barrett's oesophagus: a population-based study. *Gut* 2015;64(1):20-5. doi: 10.1136/gutjnl-2013-305506
18. Tramontano AC, Chen Y, Watson TR, et al. Esophageal cancer treatment costs by phase of care and treatment modality, 2000-2013. *Cancer Med* 2019;8(11):5158-72. doi: 10.1002/cam4.2451
19. Thein HH, Jembere N, Thavorn K, et al. Estimates and predictors of health care costs of esophageal adenocarcinoma: a population-based cohort study. *BMC Cancer* 2018;18(1):694. doi: 10.1186/s12885-018-4620-2
20. Wildi SM, Cox MH, Clark LL, et al. Assessment of health state utilities and quality of life in patients with malignant esophageal Dysphagia. *Am J Gastroenterol* 2004;99(6):1044-9. doi: 10.1111/j.1572-0241.2004.30166.x
21. Pourrahmat MM, Kim A, Kansal AR, et al. Health state utility values by cancer stage: a systematic literature review. *Eur J Health Econ* 2021;22(8):1275-88. doi: 10.1007/s10198-021-01335-8
22. Kamboj AK, Katzka DA, Iyer PG. Endoscopic Screening for Barrett's Esophagus and Esophageal Adenocarcinoma: Rationale, Candidates, and Challenges. *Gastrointest Endosc Clin N Am* 2021;31(1):27-41. doi: 10.1016/j.giec.2020.08.002
23. Peery AF, Dellon ES, Lund J, et al. Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology* 2012;143(5):1179-87 e3. doi: 10.1053/j.gastro.2012.08.002
24. Menezes A, Tierney A, Yang YX, et al. Adherence to the 2011 American Gastroenterological Association medical position statement for the diagnosis and management of Barrett's esophagus. *Dis Esophagus* 2015;28(6):538-46. doi: 10.1111/dote.12228
25. Coronado GD, Jenkins CL, Shuster E, et al. Blood-based colorectal cancer screening in an integrated health system: a randomised trial of patient adherence. *Gut* 2024;73(4):622-28. doi: 10.1136/gutjnl-2023-330980
26. Aziz Z, Wagner S, Agyekum A, et al. Cost-Effectiveness of Liquid Biopsy for Colorectal Cancer Screening in Patients Who Are Unscreened. *JAMA Netw Open* 2023;6(11):e2343392. doi: 10.1001/jamanetworkopen.2023.43392

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1: *Study design.* The overall study design was divided into three phases parts. EDRN phase I was initially based on tissue, and then transitioned to a blood-based part. The TCGA dataset provided expression-level data from 2082 miRNAs. Fourteen of these miRNAs were significantly upregulated in cases with EAC compared to controls and had a discriminatory AUC >70%. To confirm that these were differentially expressed not only in EAC but also in its precursor lesions, the expression level of these miRNAs was assessed in the GSE16456 data set after microarray normalization to confirm their upregulation in EAC, BE-HGD, and BE-LGD vs. matched normal healthy mucosa. Ten miRNAs passed these preliminary quality measures and were given full consideration as EAC-specific biomarkers. To assess whether these tissue-based results could be replicated with a different quantification method (RT-qPCR vs. sequencing vs. microarray), their expression level was internally assessed in 84 tissue samples from EAC tissue vs. normal matched mucosa, and we could confirm differential expression for nine. We then tested whether these candidates could be detected in blood. We assessed the delta CT values of these miRNAs against a normalizer miRNA (U6) in a cohort comprising 108 blood specimens and could confirm their detectability as well as their differential expression for six miRNAs, which, ultimately, constituted the foundation of the diagnostic assay. This analysis concluded phase I of our EDRN-compliant study. The final diagnostic model (termed "*EMERALD*") comprises a two-level stacked model. In the first level, two independent ML models (XGBoost and AdaBoost) are allowed to learn independently from the miRNA expression levels. In the second level, a logistic regression function is fit on these two risk scores. At the end of this process (EDRN phase II), the diagnostic model is fully locked and independently applied to a second independent cohort of 297 individuals (EDRN phase III).

Abbreviations: BE, Barrett's Esophagus; EAC, Esophageal Adenocarcinoma; HGD, Barrett's Esophagus with High-Grade Dysplasia; LGD, Barrett's Esophagus with Low-Grade Dysplasia; RT-qPCR, Reverse Transcription quantitative polymerase chain reaction; NDC, Non-disease controls.

Supplementary Figure 2: Trajectory testing for the miRNA candidates in GSE 16456 after bias correction and data standardization. Ten of the 14 miRNAs identified by differential gene expression in the TCGA dataset, which only includes EAC vs. controls, were also differentially expressed between the normal mucosa vs. LGD, HGD, and EAC (all patient-matched). These 10 miRNAs were carried over to the next phase of the study.

Supplementary Figure 3: Tissue validation for miRNA candidates in a clinical cohort of esophageal cancer patients with matching adjacent normal mucosae. Nine of the total 10 miRNA candidates discovered during the *in silico* analysis were significantly upregulated in EAC tissue samples (n = 42) compared to patient-matched adjacent normal tissues (n = 42) when measured by RT-qPCR (* $p<0.05$, ** $p<0.01$, *** $p<0.001$, ns, not significant; two-sided paired Student's t-tests).

Supplementary Figure 4: EMERALD microRNA expression in cancers other than EAC. *EMERALD* is based on six microRNAs, whose gene expression is higher in individuals with EAC/BE than non-disease controls. In order to evaluate the tissue specificity of these microRNAs, we evaluated whether their gene expression was elevated in two large publicly available datasets of patients with various cancers whose blood levels of these microRNAs were assessed in studies different than this. The expression level of the EMERALD microRNAs was separately evaluated in two independent datasets (GSE113486 and GSE113740). In the blood of individuals with cancers other than EAC/BE, compared to matched controls, no statistically significant elevation in the gene expression of any of the EMERALD microRNAs, except for hsa-18a-5p. However, no statistically significant differences for the same cancer type were observed in both cohorts simultaneously.

Supplementary Figure 5: Expression level of the six circulating miRNAs in the Italian (*training*) cohort. In the Italian cohort (N=160, 96 vs. 64), 20 had stage 0 disease (pT_{is}). The 6 miRNAs (hsa-miR-106b-5p, hsa-miR-146a-5p, hsa-miR-15a-5p, hsa-miR-18a-5p, hsa-miR-21-5p, hsa-miR-93-

5p) were all significantly upregulated in EAC/HGD compared with healthy controls in the training cohort (**** p<0.0001; two-sided Student's t-tests).

Supplementary Figure 6: Final model architecture and contribution of each miRNA to the final model structure. (**A**) *EMERALD* relies on a stacked machine-learning approach. Two different boosting-based models, XGBoost and AdaBoost, are both allowed to learn patterns from the Delta-CT values of the six miRNA candidates. At the end of their training, a logistic regression function is passed on to their results, constituting the final '*EMERALD*' model. This model is entirely developed, trained, and fine-tuned in the Italian cohort, fully locked, and then passed to the UK/Irish cohort for final external and independent testing; (**B**) Beeswarm plot of SHAP values of each microRNA for XGBoost: values further from 0 contributed more towards the accuracy of the prediction; absolute miRNA measurements are represented from lowest -yellow- to highest -purple-; (**C**) Gain is the relative contribution of each miRNA to the model, calculated by taking each miRNA's contribution for each tree. miRNAs with higher gain metrics are more important and contribute more to the overall accuracy of the model; (**D**) Beeswarm plot of SHAP values of each microRNA for AdaBoost: values further from 0 contributed more towards the accuracy of the prediction; absolute miRNA measurements are represented from lowest -yellow- to highest -purple-; (**E**) Importance of each miRNA to AdaBoost. miRNAs with a higher importance contribute more to the overall accuracy of the AdaModel.

Abbreviations: XGB, eXtreme Gradient Boosting; ADA, Adaptive Boosting; SHAP, SHapley Additive exPlanations

Supplementary Figure 7: Symptom-specific evaluation of the EMERALD blood-based test in the external and independent testing cohort. (**A**) Violin plot of *EMERALD* values according to the presence of epigastralgia-related symptoms, with and without reflux esophagitis or hiatal hernia vs. cases with BE or LGE and cases with EAC or HGD (**** p<0.0001; two-sided Student's t-tests). (**B,C**) AUROCs of the two first-level classifiers (XGB: thin line; ADA, dashed line) and the stacked model (thick line with 95% confidence intervals) in a sub-analysis of the UK/Irish *(testing*) cohort which only included symptomatic controls vs. Barrett's esophagus cases, with or without low-

grade dysplasia (**B**) and cases with either high-grade dysplasia or adenocarcinoma (**C**). (**D**) Violin plot of *EMERALD* values according to the presence of dysphagia, regurgitation, or odynophagia symptoms, with and without reflux esophagitis or hiatal hernia vs. cases with BE or LGE and cases with EAC or HGD (**** p<0.0001; two-sided Student's t-tests). (**E,F**) AUROCs of the two first-level classifiers (XGB: thin line; ADA, dashed line) and the stacked model (thick line with 95% confidence intervals) in a sub-analysis of the UK/Irish *(testing*) cohort which only included symptomatic controls vs. Barrett's esophagus cases, with or without low-grade dysplasia (**E**) and cases with either high-grade dysplasia or adenocarcinoma (**F**).

Supplementary Figure 8: Multivariate logistic regression to predict whether EMERALD-positive individuals have low-risk positive results (BE/LGD) or high-risk positive results (EAC/HGD) (results from the independent and external validation cohort). (**A**) AUROC of the ability to discriminate between EAC/HGD vs. BE/LGD among the EMERALD test-positive individuals (**B**) Confusion matrix of the EMERALD test results in the external and independent testing cohort. Sensitivity and specificity values presented outside of the confusion matrix pertain to the EMERALD blood-based test. Percentages inside the boxes represent sensitivity and specificity values of the logistic regression model in the independent and external validation cohort. 95% confidence intervals are computed according to the Wilson's method.

Supplementary Table 1. Clinical characteristics of patients and healthy participants were used for analyses in the in silico, tissue, test, training, and validation cohorts.

	Tissue (N=218)		Blood (N=565) [c]		
Total [c] N=783	***In silico* Cohorts (N = 134) [d]**	**Tissue Cohort (N = 84) [e]**	**Development Cohort (N = 108) [f]**	**Training Cohort (N = 160) [g]**	**Validation Cohort [c] (N = 297) [h,i]**
Cases [c] (N=544)	**105**	**42**	**78**	**96**	**223 [c]**
Sex					
Men	90	31	44	90	181
Women	15	11	20	6	42
Undisclosed	0	0	14	0	0
P-value	*0. 40*	*1.00*	*0.11*	*0.10*	*0.60*
Age (years)					
Median (range)	69 (27-86)	60 (35-70)	59 (28-87)	62 (27-83)	66 (34-84)
P-value	*0.56*	*1.00*	*0.57*	*0.41*	*0.44*
EAC Precursors (N=149)					
BE	0	0	20	0	88
LGD	5	0	7	0	10
HGD	5	0	7	0	7
Total	**10**	**0**	**34**	**0**	**105**
EAC (N=395)					
0 (T_{is})	0	0	0	23	0
IA	3	5	0	4	0
IB	11	4	2	11	0
IIA	7	12	3	15	0
IIB	20	3	7	10	0
IIIA	13	5	5	12	0
IIIB	8	10	5	9	0
IIIC	11	3	8	11	0
IV	11	0	5	1	0
Unavailable	11	0	9	0	118
Total	**95**	**42**	**44**	**96**	**118**
Non-disease controls (N=239) [c]	**29 [a]**	**42 [b]**	**30**	**64**	**74 [c]**
Sex					
Men	23	31	22	55	58
Women	6	11	8	9	14
Undisclosed	0	0	0	0	2
P-value	*0. 40*	*1.00*	*0.11*	*0.10*	*0.60*
Age (years)					

Median (range)	73 (45-83)	60 (35-70)	55 (35-72)	58 (35-74)	59 (38-78)
P-value	*0.56*	*1.00*	*0.57*	*0.41*	*0.44*
Symptoms, if applicable	Not applicable	Not applicable			
Persistent or treatment-refractory heartburn	--	--	0	0	47
Epigastralgia	--	--	0	0	51
Dysphagia, Odynophagia, or Regurgitation			0	0	39
Others (Sore throat, cough, globus, nausea)			0	0	9
Family history of BE			0	0	4
Not available	--	--	30	64	0

Abbreviations: BE, Barrett's esophagus; GERD, Gastro-Esophageal Reflux Disease; HGD, High-grade dysplasia; LGD, Low-grade dysplasia; RE, Reflux Esophagitis

Footnotes:

[a] Sixteen are adjacent normal esophageal tissue matched to an equal number of cases;

[b] All are adjacent normal esophageal tissue;

[c] Numbers are provided without the 9 samples that did not pass quality control;

[d] Sources: TCGA (n=89) and GSE16456 (n=32);

[e] A single-center case-only study conducted in the Netherlands in patients who had received esophagectomy (Ethical Committee approval obtained from the Radboud University Medical Centre);

[f] A multi-institutional case-control study that involved the Norton Thoracic Institute at St. Joseph's Hospital and Medical Center, USA, and the Baylor University Medical Center, USA (Ethical Committee approval obtained from both institutions separately);

[g] A single-center case-control study conducted in Italy, case-control study (Ethical Committee approval obtained from the University of Padua);

[h] The FINBAR (Factors Influencing the Barrett's Adenocarcinoma Relationship) Study was an all-Ireland population-based case-control study (n=206), all recruited from both Northern Ireland and the Republic of Ireland (Ethical committee approval obtained from the Research Ethics Committee of Queen's University, Belfast, Northern Ireland; Clinical Research Ethics Committee of the Cork Teaching Hospitals; and Research Ethics Committee Board of St. James's Hospital. Ethical approval was separately obtained from Queen's University for the recruitment of reflux esophagitis patients);

[i] A multi-institutional case-control study (n=91) that involved the Johns Hopkins Hospital, Baltimore, MD, and the Translational Genomics Research Institute, Phoenix, AZ (Ethical Committee approval obtained from both institutions separately)

Supplementary Table 2 Base-case values in cost-effectiveness modelling

	Value	Ref.
Cost		
miRNA assay	50	Estimated from internal costs
Endoscopy after a positive miRNA assay	500	Estimated from internal costs
Biopsy	150	Estimated from internal costs
Dysplasia treatment	1500	Estimated from internal costs
Staging		
Stage I	8248	[9,18-21]
Stage II	9171	
Stage III	9249	
Stage IV	9119	
Surgery		
Stage I	72914	[9,18-21]
Stage II	56124	
Stage III	63592	
Stage IV	52908	
Medical therapy (initial costs)		
Stage I	6240	[9,18-21]
Stage II	7478	
Stage III	8429	
Stage IV	9263	
Medical therapy (continuous costs)		
Stage I	2338	[9,18-21]
Stage II	2893	
Stage III	3394	
Stage IV	4334	
Follow-up visits	700	[9,18-21]
Terminal care	20599	[9,18-21]
Utility Values		
Stage 1	0.99	[9,18-21]
Stage 2	0.80	
Stage 3	0.54	
Stage4	0.52	
Cured	0.99	

Supplementary Table 3 Cost-effectiveness analysis for non-invasive screening at three different intervals, vs. invasive screening, vs. no screening

	EMERALD-based screening			Endoscopy-based screening	No screening
	Every 1 year	Every 3 years	Every 5 years		
Cohort characteristics					
Cohort size	1000	1000	1000	1000	1000
Duration of follow-up	30 years	30 years	30 years	30 years	30 years
Total number of incident cancer diagnoses	8	23	35	48	63
Total number of deaths	306	323	334	366	378
Compliance with a screening strategy	45%	45%	45%	10%	N/A
Results					
Stage at diagnosis					
Precancerous lesions intercepted	55	40	28	15	0
Stage I	1	3	5	5	3
Stage II	1	2	4	5	4
Stage III	1	2	3	6	7
Stage IV	5	16	23	42	49
Cancer-attributable deaths averted	72	55	44	12	N/A
Cost-effectiveness analysis					
Total costs (USD, M)	182.3	119.0	112.0	120.0	119.8
Total QALY	26755	26520	26382	26039	25921
Costs (and savings) (USD, M)	62.5	(0.8)	(7.8)	0.2	*Ref.*
Effect difference	834	599	461	748	*Ref.*
Costs (and savings) per additional QALY (USD)	74940.05	(1335.56)	(16919.74)	267.38	*Ref.*

Abbreviations: M, Millions; QALY, Quality-adjusted life years; USD, US dollars