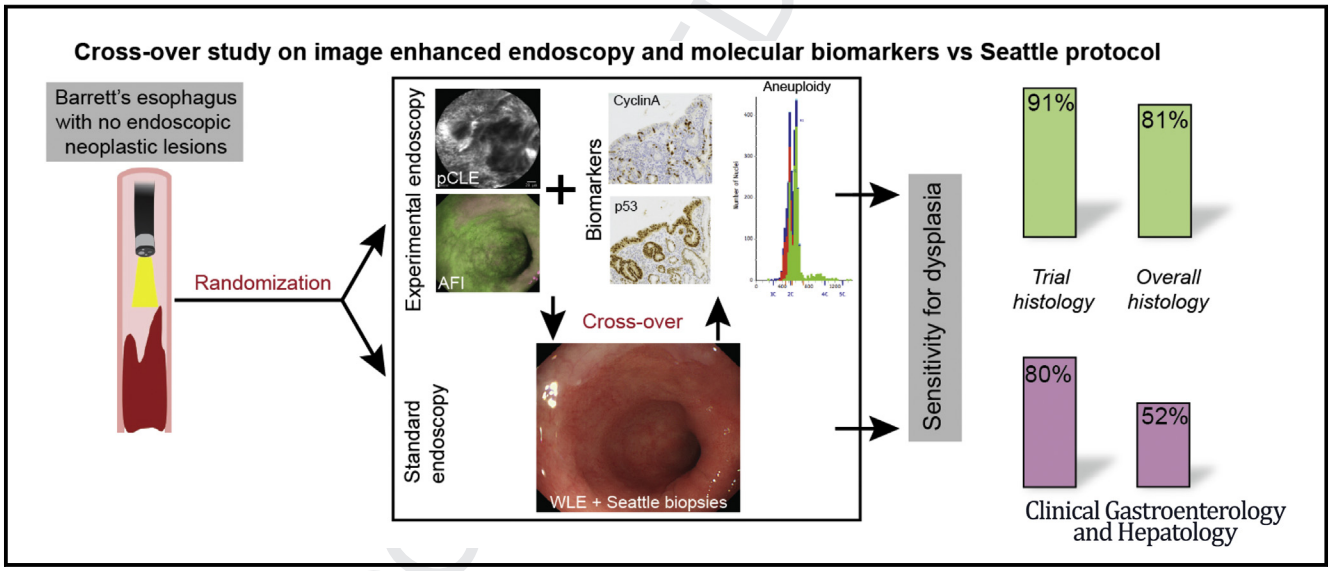
Image Enhanced Endoscopy and Molecular Biomarkers Vs Seattle Protocol to Diagnose Dysplasia in Barrett's Esophagus

Mathew Vithayathil,[*] Ines Modolell,[‡] Jacobo Ortiz-Fernandez-Sordo,[§] Dahmane Oukrif,[||] Apostolos Pappas,[*] Wladyslaw Januszewicz,[*,¶] Maria O'Donovan,[#] Andreas Hadjinicolaou,[*,‡] Michele Bianchi,[*] Adrienn Blasko,[*] Jonathan White,[§] Philip Kaye,[**] Marco Novelli,[||] Lorenz Wernisch,[‡‡,§§] Krish Ragunath,[§] and Massimiliano di Pietro[*,‡]

*MRC Cancer Unit, [§§]MRC Biostatistics Unit, University of Cambridge, United Kingdom; [‡]Department of Gastroenterology, [#]Department of Histopathology, Cambridge University Hospital NHS Foundation Trust, United Kingdom; [§]Nottingham Digestive Diseases Centre, NIHR Nottingham Biomedical Research Centre, **Department of Histopathology, Nottingham University Hospitals NHS Trust, University of Nottingham, United Kingdom; [||]Department of Histopathology, University College London Hospital, Longdon, United Kingdom; [¶]Department of Gastroenterology, Hepatology and Clinical Oncology, Medical Centre for Postgraduate Education, Warsaw, Poland; [‡‡]BIOS Health, Ltd, Cambridge, United Kingdom



BACKGROUND & AIMS: Dysplasia in Barrett's esophagus often is invisible on high-resolution white-light endoscopy (HRWLE). We compared the diagnostic accuracy for inconspicuous dysplasia of the combination of autofluorescence imaging (AFI)-guided probe-based confocal laser endomicroscopy (pCLE) and molecular biomarkers vs HRWLE with Seattle protocol biopsies.

METHODS: Barrett's esophagus patients with no dysplastic lesions were block-randomized to standard endoscopy (HRWLE with the Seattle protocol) or AFI-guided pCLE with targeted biopsies for molecular biomarkers (p53 and cyclin A by immunohistochemistry; aneuploidy by image cytometry), with crossover to the other arm after 6 to 12 weeks. The histologic end point was a diagnosis from all study biopsies (trial histology). A sensitivity analysis was performed for

Abbreviations used in this paper: AFI, autofluorescence imaging; BE, Barrett's esophagus; CLE, confocal laser endomicroscopy; GI, gastrointestinal; HGD, high-grade dysplasia; HRWLE, high-resolution white-light endoscopy; IMC, intramucosal carcinoma; LGD, low-grade dysplasia; NBI, _____; NDBE, nondysplastic Barrett's esophagus; pCLE, probe-based confocal laser endomicroscopy.

1542-3565
https://doi.org/10.1016/j.cgh.2022.01.060

overall histology, which included diagnoses within 12 months from the first study endoscopy. Endoscopists were blinded to the referral endoscopy and histology results. The primary outcome was diagnostic accuracy for dysplasia by real-time pCLE vs HRWLE biopsies.

RESULTS: **Of 154 patients recruited, 134 completed both arms. In the primary outcome analysis (trial histology analysis), AFI-guided pCLE had similar sensitivity for dysplasia compared with standard endoscopy (74.3%; 95% CI, 56.7–87.5 vs 80.0%; 95% CI, 63.1–91.6; $P = .48$). Multivariate logistic regression showed pCLE optical dysplasia, aberrant p53, and aneuploidy had the strongest correlation with dysplasia (secondary outcome). This 3-biomarker panel had higher sensitivity for any grade of dysplasia than the Seattle protocol (81.5% vs 51.9%; $P < .001$) in the overall histology analysis, but not in the trial histology analysis (91.4% vs 80.0%; $P = .16$), with an area under the receiver operating curve of 0.83.**

CONCLUSIONS: **Seattle protocol biopsies miss dysplasia in approximately half of patients with inconspicuous neoplasia. AFI-guided pCLE has similar accuracy to the current gold standard. The addition of molecular biomarkers could improve diagnostic accuracy.**

Keywords: Barrett's Esophagus; Esophageal Adenocarcinoma; Dysplasia; Confocal Laser Endomicroscopy; Autofluorescence.

Barrett's esophagus (BE) is the only known precursor lesion to esophageal adenocarcinoma.[1] BE has an estimated risk of progression to cancer of 0.3% per year, which increases 10- to 50-fold when low-grade dysplasia (LGD) and high-grade dysplasia (HGD) are diagnosed.[2–4] Treatment of dysplastic BE with endoscopic ablation prevents progression to cancer,[4,5] therefore endoscopic surveillance of BE is recommended.[6,7] Because dysplasia can be invisible on high-resolution white-light endoscopy (HRWLE), nontargeted biopsies are recommended according to the Seattle protocol.[6,7] However, adherence to this protocol is poor in clinical practice because it is laborious and time consuming.[8] In addition, interobserver agreement among histopathologists for a dysplasia diagnosis is suboptimal.[3,9] Finally, random biopsies can miss inconspicuous dysplasia. To date, there are scarce data on the true sensitivity of Seattle protocol biopsies in patients without endoscopically visible lesions.

Confocal laser endomicroscopy (CLE) provides real-time microscopic visualization of gastrointestinal mucosa. CLE diagnostic criteria for LGD and HGD in BE have been established.[10,11] Sharma et al[12] showed that the combination of HRWLE, NBI, and CLE achieved a sensitivity of 100% and a specificity of 55.7% for HGD and intramucosal carcinoma (IMC). Similarly, Canto et al[13] showed the addition of CLE to HRWLE increased sensitivity for Barrett's neoplasia from 40% to 96%. These trials included patients with flat BE and mucosal lesions suspicious of early neoplasia, which can influence the pretest endomicroscopic diagnosis. To date, no studies have assessed the diagnostic accuracy of CLE for dysplasia in patient cohorts with inconspicuous BE only.

A limitation of CLE is the narrow field of view. Autofluorescence imaging (AFI) detects the different fluorescence properties of early BE-related neoplasia and has high sensitivity for HGD, but also a significant false-positive rate.[14] We previously showed that an AFI-positive signal in BE correlates with molecular aberrations regardless of dysplasia, suggesting that a proportion of false positivity is the result of sampling bias.[15] A 3-biomarker panel including aneuploidy, cyclin A, and p53 on AFI-targeted biopsies had sensitivity and specificity for HGD/IMC of 96% and 89%, respectively. In a feasibility study combining probe-based CLE (pCLE) with AFI, this multimodal approach achieved 96.4% sensitivity and 74.1% specificity for a diagnosis of BE-related neoplasia.[16]

We conducted a multicenter randomized crossover study with the primary aim to evaluate the diagnostic accuracy for dysplasia of AFI-guided pCLE compared with HRWLE and Seattle protocol biopsies in patients with BE and no endoscopically visible lesions. We also evaluated the added diagnostic value of molecular biomarkers, the time to perform standard and experimental procedures, and the acceptability by patients of optical dysplasia diagnosis.

Methods

Study Design

This was a prospective randomized crossover study across 2 tertiary referral centers. The study was approved by the Cambridgeshire Research Ethics Committee (09/H0308/118). Patients were block-randomized using computer-generated randomization in blocks of 4 (www.randomization.com) to receive either HRWLE with Seattle protocol biopsies (standard arm) or endoscopy with AFI-directed pCLE and targeted biopsies for molecular biomarkers (experimental arm). Patients crossed over to the other arm after 6 to 12

weeks. Different endoscopists performed procedures in the 2 arms. Endoscopists could not be blinded to the intervention arm but were blinded to the endoscopy and histology results of the pretrial endoscopy and other study arm.

Participants

Inclusion criteria were as follows: patients aged older than 18 years diagnosed with BE greater than C2 and/or M3 on pretrial endoscopy (as per the Prague Classification[17]) referred for surveillance of nondysplastic BE (NDBE) or assessment of flat dysplasia. The reason for inclusions of BE segments at least C2 or M3 was 2-fold: image-enhanced assisted detection is expected to be more advantageous for long-segment BE, and AFI has a high false-positive rate at the esophagogastric junction.[15] Exclusion criteria were as follows: previous evidence of BE-related neoplasia visible on endoscopy, previous histologic evidence of esophageal adenocarcinoma, esophagitis (Los Angeles grade ≥B), previous esophagectomy, fluorescein allergy, severe/uncontrolled asthma, coagulopathy or anticoagulant/antiplatelet therapy for high-risk conditions, active/severe cardiopulmonary disease, or decompensated liver disease.

Study Outcomes

The primary outcome was the diagnostic accuracy for dysplasia of AFI-guided pCLE using the trial histology as the gold standard. Secondary outcomes included the following: (1) diagnostic accuracy of AFI-guided pCLE for dysplasia with reference to the overall histology, which included biopsy specimens taken within 12 months before enrollment; (2) added diagnostic value of molecular biomarkers; (3) time to perform the endoscopy; and (4) patient-reported experience related to experimental and standard endoscopy.

Endoscopic Procedures

Patients received 2 endoscopic procedures within the trial duration. In the standard arm, HRWLE only was allowed for inspection using FQ260Z, HQ290, or H290Z endoscopes (Olympus, Tokyo, Japan). Subtle lesions were allowed if not clearly in keeping with BE-related neoplasia, and therefore received targeted biopsies. Random biopsy specimens then were taken every 2 cm of the length of BE. In the experimental arm, FQ260Z endoscopes were used. The initial inspection was performed with HRWLE only. The endoscopist then switched to AFI mode and areas of purple–red color within a green background (AFI+) were identified (Figure 1). At the discretion of the endoscopists, AFI+ lesions were marked with argon-plasma coagulation (VIO 200; ERBE, Tuebingen, Germany) or snare tip to delineate the area of interest. AFI+ areas, together with subtle HRWLE lesions if present, then were studied with pCLE after intravenous fluorescein (10% solution, 2.5 mL) and then received 2 targeted biopsies stored in formalin. At least 2 pCLE videos per endoscopic location were recorded. A maximum of 4 AFI+ areas per patient were allowed for pCLE analysis. In patients with no AFI+ areas, 1 random location was used for pCLE analysis and targeted biopsies for every 5 cm of BE maximum extent. The endoscopist made a live pCLE diagnosis and then reviewed pCLE videos offline to make the final pCLE diagnosis. Patients with evidence of lesions at the first endoscopy that were unequivocally in keeping with BE-related neoplasia on HRWLE were excluded from the study.

What You Need to Know

Background

Endoscopic diagnosis of flat dysplasia in Barrett's esophagus is challenging. Previous trials investigating image-enhanced endoscopy (IEE) have included patients with lesions visible on white-light endoscopy.

Findings

White-light endoscopy with Seattle protocol biopsies underdiagnoses approximately half of the patients with flat dysplasia and IEE does not improve the diagnostic accuracy. Biomarkers on biopsies directed by IEE can improve the sensitivity for dysplasia.

Impact for patient care

This trial provides a methodologic model for future studies investigating the endoscopic diagnosis of flat dysplasia. Biomarkers should be used in the assessment of patients with Barrett's esophagus to inform clinical decisions.

Optical Probe-Based Confocal Laser Endomicroscopy Diagnosis

Before the study, the endoscopist received online and live pCLE training. Endoscopists reported a pCLE diagnosis at the time of endoscopy as one of the following: NDBE, LGD, or HGD. For the primary and secondary outcomes, pCLE diagnoses of LGD and HGD were regarded as dysplasia because interobserver agreement between LGD and HGD on pCLE was shown to be low.[10] Details on pCLE training and diagnostic criteria are provided in the Supplementary Materials.

Biopsy and Histology

Tissue biopsy specimens from both arms were formalin-fixed and paraffin-embedded for histopathologic assessment. Biopsy specimens were reviewed by a

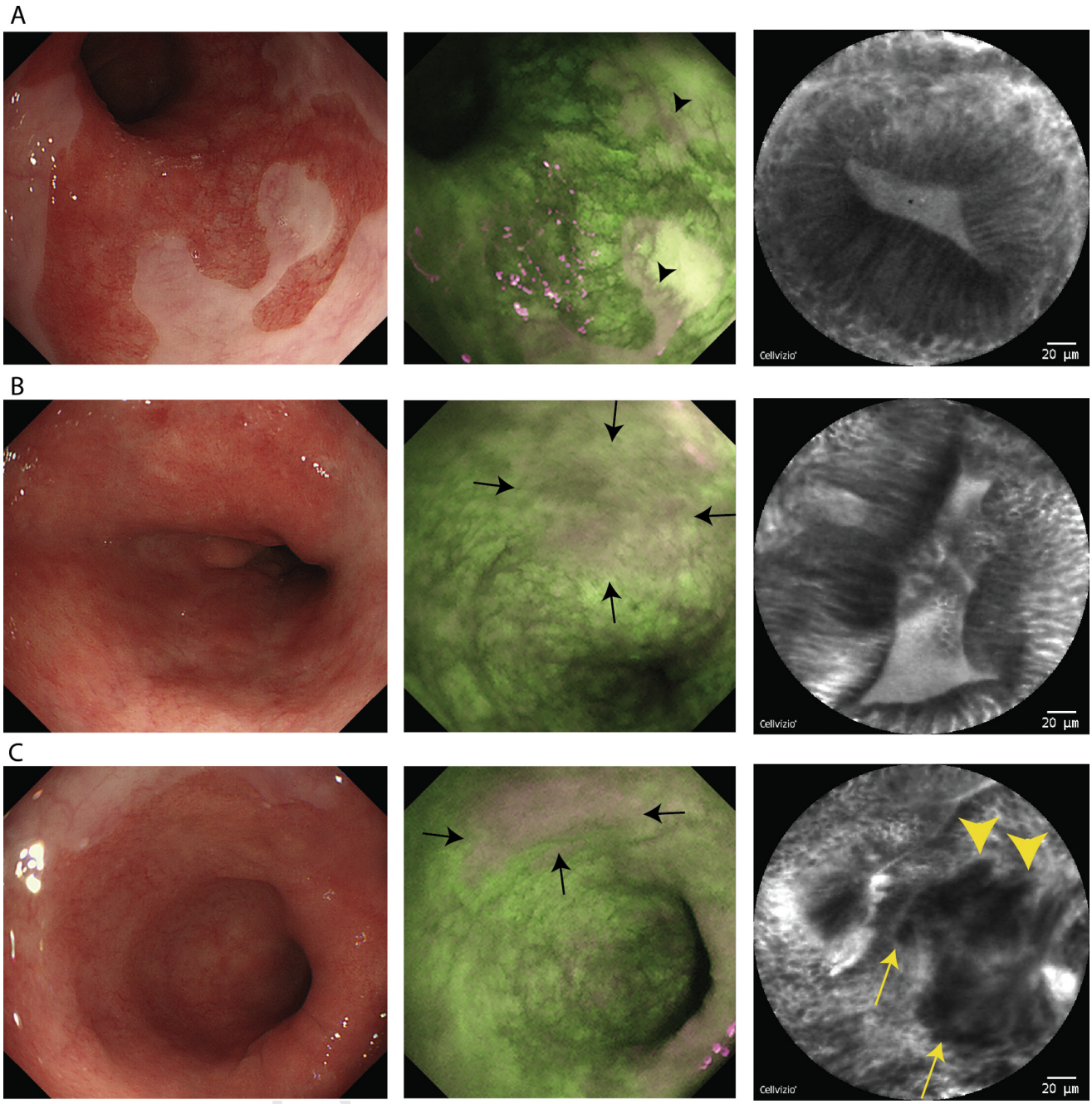


Figure 1. Examples of image-enhanced endoscopic diagnosis. (*A*) Nondysplastic Barrett's esophagus (NDBE) with negative imaging features. Flat inconspicuous BE on high-resolution white-light endoscopy (HRWLE) (*left*). Nonsuspicious green signal on autofluorescence imaging (AFI) (*middle*); *arrowheads* indicate false-positive AFI signal at the level of the squamous columnar junction. Probe-based confocal laser endomicroscopy (pCLE) on random location showed nondysplastic glands with regular contours and epithelial cells with regular columnar shape (*right*). (*B*) NDBE with AFI+ signal. Flat inconspicuous BE on HRWLE (*left*). AFI+ area at the 12 o'clock position (*middle*, *arrows*). pCLE on targeted location (*right*) showed nondysplastic glands with smooth margins and uniform columnar cells but no obvious goblet cells. (*C*) Example of AFI+ BE with high-grade dysplasia (HGD). HRWLE shows featureless BE with possibly subtle pale discoloration at the 12 o'clock position (*left*). AFI showed clear positive signal (*middle*, *arrows*). pCLE showed dysplastic glands (*right*) with irregular shape, saw-toothed margins (*arrowheads*), and cellular pleomorphism (*arrows*).

gastrointestinal (GI) pathologist with extensive expertise in BE in accordance with the Vienna classification.[18] All dysplastic cases, including indefinite dysplasia were reviewed by a second expert GI pathologist from the other institution, with consensus diagnosis achieved for discordant cases. For the purpose of the analysis, indefinite dysplasia was grouped with NDBE. In the standard arm, p53 immunohistochemistry was performed at the discretion of the pathologist, as per the standard of care.

Procedural Time

The time taken to perform each arm of the trial was recorded. The start time was the time of insertion of the endoscope and the end time was the time of the patient's extubation.

Molecular Biomarker Assays

A 3-biomarker panel including cyclin A, p53, and aneuploidy was selected based on previously published data.[15,16] Cyclin A and p53 expression were assessed with immunohistochemistry and aneuploidy with image cytometry. A full panel of biomarkers was available in 96.3% of cases. Details on biomarker methodology are provided in the Supplementary Materials.

Statistical Analysis

In per-lesion analysis, the sensitivity and specificity for dysplasia of pCLE and HRWLE (presence vs absence of mucosal lesion) were calculated in reference to the histologic diagnosis at each AFI-targeted area. In per-patient analysis, the gold standard diagnosis was the highest grade of dysplasia detected on biopsy specimens from both arms (trial histology). Diagnostic accuracy was calculated for the Seattle protocol (sensitivity) and pCLE diagnosis (sensitivity and specificity). The McNemar test compared differences between the Seattle protocol and pCLE diagnosis. A sensitivity analysis was performed using the combination of trial histology and any histology from endoscopies performed up to 12 months before enrollment in the trial (overall histology) as reference. All cases of pretrial histology were reviewed by the trial GI pathologists.

The diagnostic accuracy for the addition of molecular biomarkers was determined. Multivariate logistic regression including optical dysplasia by pCLE, p53 expression, cyclin A expression, and aneuploidy was performed to identify the biomarkers with the strongest correlation with dysplasia. The area under the receiver operating curve was used to assess the diagnostic accuracy of the biomarker panel with different cut-off levels. A time comparison between the experimental and standard arms was performed using a paired *t* test. All authors had access to the study data and approved the final manuscript.

Sample Size

A large multicenter study showed that the Seattle protocol has a sensitivity for any grade of dysplasia of 84.6%.[19] A recent single-center study showed that AFI-targeted pCLE had a sensitivity and specificity for any grade of dysplasia of 96% and 86%, respectively.[16] With this level of diagnostic accuracy, we calculated that 47 patients with a previous diagnosis of dysplasia and 86 patients with NDBE (total, 133 patients) were required to show a sensitivity of at least 0.80 and a specificity of at least 0.75 for AFI-targeted pCLE at a significance level of 0.01. Assuming the true sensitivity may have been overestimated as a result of the small sample size in the second study, we assumed that 133 patients still would show a sensitivity of at least 0.80 and a specificity of at least 0.75 at a significance level of 0.05. Considering a potential dropout of 10% after the first endoscopy, the prespecified sample size was 146.

Results

A total of 154 patients were recruited between May 2017 and October 2019, of whom 8 were excluded based on first endoscopy findings (macroscopic lesions clearly in keeping with BE-related neoplasia, short segment of BE, or esophagitis). One patient was excluded because of a protocol breach (acetic acid chromoendoscopy on standard-arm endoscopy) and 11 patients withdrew consent before the second endoscopy. As shown in Figure 2, there were 134 patients who completed both arms of the study. Patient characteristics are shown in Table 1. Eighteen patients (13.4%) had a trial histologic diagnosis of HGD/IMC, while 17 (12.7%) were diagnosed with LGD. The HGD/IMC diagnosis was made in 4 cases in the experimental arm only, 5 cases in the standard arm only, and in 8 cases in both arms. Any grade of dysplasia was found in 7 cases in the experimental arm only, 14 in the standard arm only, and 14 in both arms.

AFI had a sensitivity for dysplasia of 88.9%, but a false-positive rate greater than 80%. In 18.1% (n = 41) of these areas the endoscopist noticed a subtle abnormality on HRWLE, however, the patients were retained in the study because the endoscopist did not judge the lesion unequivocally neoplastic; HGD or LGD was confirmed in 12.2% (n = 5) and 9.8% (n = 4) of these subtle lesions, respectively. Of the 278 targeted areas, 28.8% showed optical dysplasia on pCLE (n = 80). In the standard arm, 67 patients (50%) had subtle mucosal irregularity and received targeted biopsies for a total of 116 endoscopic areas. Of these areas, 10.3% (n = 12) showed HGD/IMC and 6.9% (n = 8) showed LGD. Targeted biopsies from the standard arm identified dysplasia in 12.7% of patients (HGD/IMC, n = 10; LGD, n = 7).

Performance of Endoscopic Techniques

In the per-lesion analysis, the optical diagnosis by pCLE had a sensitivity and specificity for HGD/IMC of 69.2% (95% CI, 38.6–90.9) and 73.2% (95% CI, 67.5–78.4) respectively, and for any grade of dysplasia of 66.7% (95% CI, 46.0-83.5) and 75.3% (95% CI, 69.5-80.5), respectively. Within the experimental arm, pCLE had a higher sensitivity than HRWLE for HGD/IMC (P = .046) and all grades of dysplasia (P = .01), but lower specificity (HGD/IMC, P = .01; all grades of dysplasia,

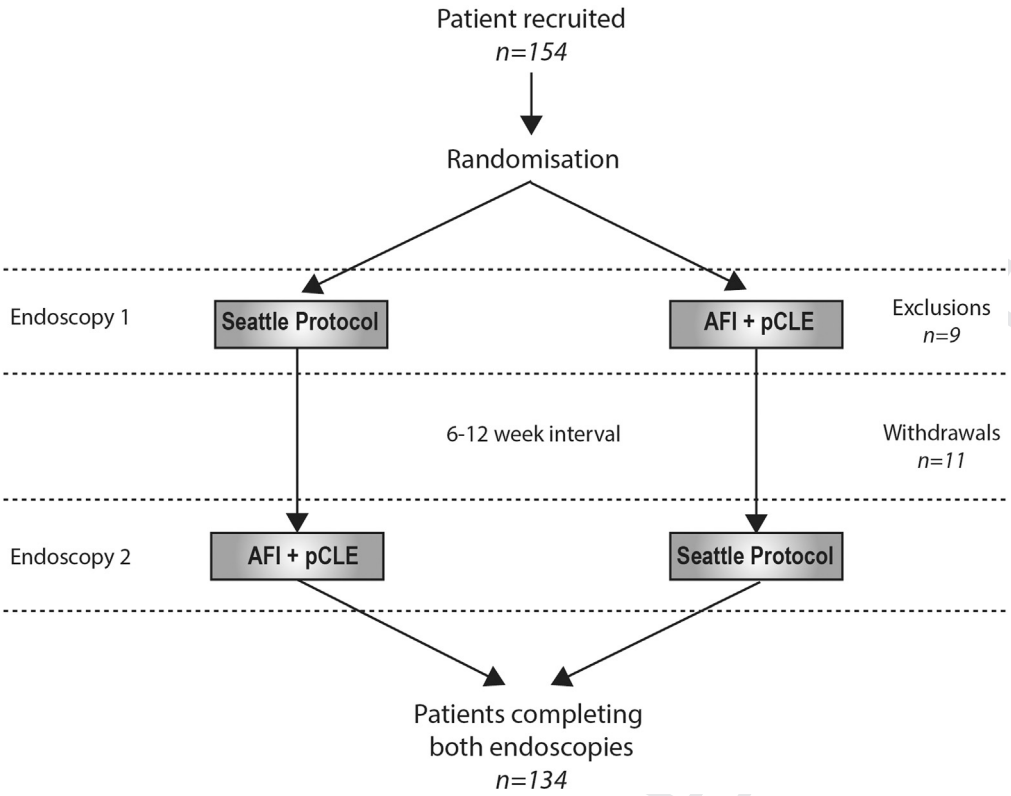


Figure 2. Recruitment and cross-over study flowchart. AFI, autofluorescence imaging; pCLE, probe-based confocal laser endomicroscopy.

$P = .02$) (Supplementary Table 1). The agreement between the live pCLE diagnosis and the offline pCLE diagnosis was substantial (K = 0.76; SE, 0.04).

In per-patient analysis, there was no difference in the sensitivity of pCLE for dysplasia compared with HRWLE with the Seattle protocol (76.5%; 95% CI, 50.1–93.2 vs 76.5%; 95% CI, 50.1–93.2, respectively; $P = 1.00$ for HGD/IMC; 74.3%; 95% CI, 56.7–87.5 vs 80.0%; 95% CI, 63.1–91.6, respectively; $P = .48$, for all grades of dysplasia) (Table 2). The use of AFI-targeted pCLE led to 2.1 optical biopsy specimens per patient on average compared with 12.3 tissue biopsy specimens taken in the Seattle protocol.

Sampling error is a well-known limitation of the Seattle protocol. To capture cases of dysplasia missed in the trial, we performed a sensitivity analysis including pretrial histology (overall histology). Overall, 54 patients had dysplasia of any grade, with 13 and 6 additional cases of HGD and LGD, respectively (Table 3). Standard endoscopy missed 28 cases of dysplasia (miss rate, 51.9%), 11 of which were detected by experimental endoscopy. Experimental endoscopy missed 20 dysplastic cases (miss rate, 37%), of which 5 were diagnosed correctly by standard endoscopy. In the overall histology analysis, AFI-guided pCLE had a higher sensitivity for HGD/IMC than Seattle protocol biopsies (73.3%; 95% CI, 54.1–87.7 vs 43.3%; 95% CI, 25.5–62.6, respectively; $P = .02$). The difference in sensitivity for all grades of dysplasia was not statistically significant (63.0%; 95% CI, 48.7–75.7 vs 51.9%; 95% CI, 37.8–65.7, respectively; $P = .13$). The diagnostic accuracy of AFI-targeted pCLE varied across individual operators. Two endoscopists achieved a sensitivity greater than 90%, while 2 endoscopists showed a sensitivity of less than 60% (Supplementary Table 2).

Molecular Biomarkers

In the per-patient analysis the sensitivity and specificity for dysplasia of individual biomarkers were 48.6% and 93.9% for p53, 47.1% and 69.4% for cyclin A, and 40.0% and 88.5% for aneuploidy, respectively. We performed a multivariate logistic regression analysis to identify the biomarkers with the strongest correlation with the dysplasia status, including optical dysplasia by pCLE. The model showed that p53, aneuploidy, and optical dysplasia correlated significantly with a diagnosis of dysplasia (Supplementary Table 3). A panel comprising these 3 biomarkers showed an area under the receiver operating curve of 0.83 (95% CI, 0.76–0.91) for a diagnosis of any grade of dysplasia and 0.88 (95% CI, 0.78–0.97) for a diagnosis of HGD/IMC. Using a threshold of 1 positive biomarker, this panel had a higher sensitivity than the Seattle protocol in detecting dysplasia in the overall histology analysis (81.5% vs 51.9%; $P < .001$) (Tables 2 and 3). The difference was not statistically significant in the trial histology analysis (91.4% vs 80.0%; $P = .16$).

Procedural Time

The mean time for experimental endoscopy was significantly longer than the standard endoscopy (22.3 vs

Table 1. Patient Demographics

Variable	
Total number of patients, *n*	134
Demographics	
Age, *y*	67.3 (38.0–89.0)
Male	104 (77.6)
Maximal length of Barrett's	5.9 (3.0–16.0)
Overall histologic diagnosis	
NDBE	92 (68.7)
ID	7 (5.2)
LGD	18 (13.4)
HGD/IMC	17 (12.7)
Endoscopic: experimental arm	
Total number of areas studied	278
AFI-positive areas	226 (81.3)
AFI-positive areas per patient	1.69 (1.0–4.0)
AFI-positive areas visible on HRWLE	41 (18.1)
Total number of pCLE optical biopsies	278
Optical biopsies per patient	2.1 (1.0–4.0)
Areas with pCLE dysplasia	80 (28.8)
Areas with pCLE dysplasia visible on HRWLE	27 (9.7)
Endoscopic: standard arm	
Total number of tissue blocks	520
Biopsies from targeted areas	116 (22.3)
Tissue blocks from random biopsies	404 (77.7)
Total number of tissue biopsies	1656
Tissue biopsies per patient	12.4 (2.0–33.0)

NOTE. The mean (range) for continuous variables, and n (%) for discrete variables are shown.
AFI, autofluorescence imaging; HGD, high-grade dysplasia; HRWLE, high-resolution white-light endoscopy; ID, indeterminate for dysplasia; IMC, intramucosal carcinoma; LGD, low-grade dysplasia; NDBE, nondysplastic Barrett's esophagus; pCLE, probe-based confocal laser endomicroscopy.

16.4 min; $P < .001$) (Figure 3). Because 3 of 5 endoscopists had no experience with pCLE before the trial, to assess whether the time to perform pCLE imaging improved with experience, we looked at the trend of procedural time in each quarter of the study period (quarter 1 to quarter 4). We found evidence of a learning curve, with a mean time for the experimental endoscopy decreasing during the study time (quarter 1 vs quarter 4, 26.5 vs 19.0 min, respectively; $P < .001$).

Patient Acceptability

We found that communication of the optical dysplasia diagnosis immediately after the procedure did not significantly alter anxiety levels compared with the routine standard of waiting for a histologic diagnosis. Details of patient-reported experiences are provided in the Supplementary Material and in Supplementary Figure 1.

Discussion

In this trial we found that in patients with inconspicuous BE, AFI-guided pCLE has similar diagnostic accuracy for dysplasia compared with standard HRWLE with Seattle protocol biopsies. The addition of molecular biomarkers improved the diagnostic accuracy compared with the current gold standard.

We previously generated and validated pCLE diagnostic criteria for LGD, which, in a retrospective study, diagnosed dysplasia with 82% sensitivity and 75% specificity. In this study, we validated the use of pCLE for detection of all grades of dysplasia in real time. Two randomized trials have assessed the diagnostic accuracy of CLE for BE-related dysplasia, with different designs. Sharma et al[12] investigated 101 patients with BE with a single endoscopic procedure in which HRWLE, NBI, and pCLE were used sequentially. Because 25% of the study population had a cancer diagnosis, the pretest probability in this trial was high and only 1 patient with HGD/IMC was missed by the combination of HRWLE and NBI. The study by Canto et al[13] randomized 192 patients with BE of less than 10 cm to either HRWLE or HRWLE with CLE. CLE was performed on targeted as well as random locations and the histologic end point was HGD/IMC.

The reason for using a flagging technique was to reduce the number of locations for pCLE assessment because following a Seattle protocol distribution would be time consuming. AFI was chosen based on previous evidence of feasibility and evidence that AFI-positive signal correlates with molecular aberrations.[20] We did

Table 2. Per-Patient Analysis of Diagnostic Accuracy of Seattle Protocol Histology, AFI-Targeted Optical pCLE Diagnosis, and 3-Biomarker Panel

	Seattle protocol	AFI + pCLE	*P* value[a]	3-biomarker panel	*P* value[b]	*P* value[c]
Dysplasia (all grades) (n = 35)						
Sensitivity, %	80.0	74.3	.48	91.4	.16	.01
Specificity, %	–	66.7	–	56.6	–	.002
High-grade dysplasia (n = 17)						
Sensitivity, %	76.5	76.5	1.00	94.1	.18	.083
Specificity, %	–	60.7	–	49.6	–	<.001

NOTE. The 3-biomarker panel consisted of 1 or more of optical dysplasia on pCLE, aberrant p53 on immunohistochemistry, and/or aneuploidy on flow cytometry.
AFI, autofluorescence imaging; pCLE, probe-based confocal laser endomicroscopy.
[a]*P* value calculated for McNamar test for Seattle protocol vs AFI-targeted pCLE
[b]*P* value calculated for McNamar test for Seattle protocol vs 3-biomarker panel
[c]*P* value calculated for McNamar test for AFI-targeted pCLE vs 3-biomarker panel.

Table 3. Per-Patient Analysis Including Pretrial Endoscopy of Diagnostic Accuracy of Seattle Protocol Histology, AFI-Targeted Optical pCLE Diagnosis, and 3-Biomarker Panel

	Seattle protocol	AFI + pCLE	*P* value[a]	3-biomarker panel	*P* value[b]	*P* value[c]
Dysplasia (all grades) (n = 54)						
Sensitivity	51.9	63.0	.13	81.5	<.001	.002
Specificity	–	68.8		61.3	–	.014
High-grade dysplasia (n = 30)						
Sensitivity	43.3	73.3	.02	86.7	<.001	.046
Specificity	–	64.4		52.9	–	<.001

NOTE. The 3-biomarker panel consisted of 1 or more of optical dysplasia on pCLE, aberrant p53 on immunohistochemistry, and/or aneuploidy on flow cytometry.
AFI, autofluorescence imaging; pCLE, probe-based confocal laser endomicroscopy.
[a]*P* value calculated for McNamar test for Seattle protocol vs AFI-targeted pCLE.
[b]*P* value calculated for McNamar test for Seattle protocol vs 3-biomarker paneld
[c]*P* value calculated for McNamar test for AFI-targeted pCLE vs 3-biomarker panel.

not opt for acetic acid because this alters endomicroscopic features of BE, and NBI lacks evidence for detection of LGD. However, AFI is not widely available and therefore is unlikely to be the ideal flagging technique for future applications. In the future, other imaging modalities will need to be investigated in combination with pCLE.

This randomized trial provides definitive evidence that Seattle protocol biopsies have low sensitivity for dysplasia in patients with inconspicuous BE even in expert centers. The results indicate that dysplasia can be missed in up to 50% of patients referred with early BE neoplasia and no macroscopically visible lesions. This supports the recommendation that a HGD diagnosis should prompt an ablation strategy in the appropriate patient setting when corroborated by a second pathologist regardless of whether it is confirmed at subsequent endoscopy. Likewise, given the significant sampling error, patients with LGD should be followed up with intensive surveillance even if LGD is not confirmed at immediate subsequent endoscopies. These results also provide an important comparator to gauge the utility of pan-esophageal nonendoscopic cell collection devices, such as Cytosponge, for future use in BE surveillance settings.[21]

In this study, the addition of molecular biomarkers improved the diagnostic accuracy for dysplasia in the overall histology analysis. The difference was not significant in the trial histology analysis, likely owing to the smaller number of dysplastic cases when we excluded pathology results from the endoscopy before trial procedures. Our group previously showed that p53 and aneuploidy have the best performance in identifying dysplasia and predicting future progression.[15,22] In a more recent study, aberrant p53 was associated with a hazard ratio for progression of 5.03 (95% CI, 3.88–6.5) in patients with NDBE.[23] In this study, we used only biomarkers compatible with routine clinical biopsies and used image cytometry on paraffin-embedded biopsy specimens to measure aneuploidy. In addition, p53 immunohistochemistry is used routinely as a diagnostic adjunct in many pathology laboratories. This study suggests that it is possible to achieve high diagnostic accuracy with a biomarker-aided diagnosis on biopsies targeted by optical imaging, dispensing random sampling. Future guidelines should address the role of p53

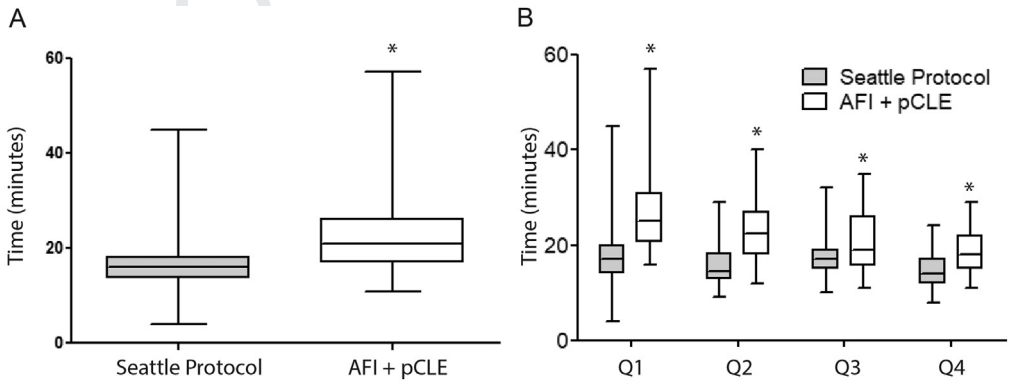


Figure 3. Time taken to perform experimental endoscopy (autofluorescence imaging [AFI]-guided probe-based confocal laser endomicroscopy [pCLE]) and standard endoscopy (high-resolution white-light endoscopy with Seattle protocol biopsies). The *boxes* represent the median and interquartile ranges, with *vertical line* ranges from minimum to maximum values. The paired Student *t* test compared paired endoscopy times from each arm. *$P < .05$. (*A*) Data for the overall cohort. (*B*) Patients divided into temporal quartiles of the study (quarter 1 to quarter 4).

and other biomarkers for risk stratification to inform clinical decisions.

We believe that this study provides an important model for the design of future endoscopy trials that aim to investigate diagnostic accuracy for inconspicuous dysplasia. First, we included only patients referred without visible lesions or, at most, with subtle visible areas of uncertain significance, which represent the most challenging group of patients. Although 50% of patients did have subtle lesions on HRWLE, there is evidence that the majority of HRWLE visible lesions are indeed NDBE with a positive predictive value varying between 27% and 42%.[12] In this trial, only 22% of these subtle lesions harbored dysplasia. Second, we used all grades of dysplasia as the histologic end point in a prospective randomized trial. The majority of endoscopy trials for BE-related neoplasia focused on detection of HGD/IMC. However, LGD carries a significant risk of progression to HGD/IMC of up to 10% per year,[3] which can be reduced significantly by endoscopic ablation.[5] Finally, the crossover design allowed a direct comparison between the gold standard and experimental imaging within the same patient.

Our study had limitations. First, referral histology within the prior 12 months was only available in 64.2% of cases. Second, because of the crossover design we could not exclude that prior biopsy sites may have appeared as mucosal irregularities on a second endoscopy. We found variations in performance in experimental endoscopy, with 2 operators having a low sensitivity for detecting dysplasia. Finally, the study was performed in 2 high-volume tertiary referral centers, therefore the results might not be applicable to a general endoscopy service.

In conclusion, this study confirms and quantifies the low sensitivity of the Seattle protocol for inconspicuous dysplasia. Although it is possible to achieve a similar level of diagnostic accuracy with image-enhanced endoscopy, challenges related to the duration of the endoscopy with complex endoscopy protocols remain. Molecular biomarkers can improve diagnostic accuracy and should be implemented into clinical practice.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Clinical Gastroenterology and Hepatology* at www.cghjournal.org, and at https://doi.org/10.1016/j.cgh.2022.01.060

References

1. Pohl H, Sirovich B, Welch HG. Esophageal adenocarcinoma incidence: are we reaching the peak? Cancer Epidemiol Biomarkers Prev 2010.
2. Desai TK, Krishnan K, Samala N, et al. The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. Gut 2012.
3. Duits LC, Phoa KN, Curvers WL, et al. Barrett's oesophagus patients with low-grade dysplasia can be accurately risk-stratified after histological review by an expert pathology panel. Gut 2015.
4. Shaheen NJ, Sharma P, Overholt BF, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. N Engl J Med 2009;360:2277–2288.
5. Phoa KN, van Vilsteren FGI, Weusten BLAM, et al. Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia. JAMA 2014.
6. 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.
7. Shaheen NJ, Falk GW, Iyer PG, et al. ACG clinical guideline: diagnosis and management of Barrett's esophagus. Am J Gastroenterol 2016.
8. Abrams JA, Kapel RC, Lindberg GM, et al. Adherence to biopsy guidelines for Barrett's esophagus surveillance in the community setting in the United States. Clin Gastroenterol Hepatol 2009.
9. Wani S, Falk GW, Post J, et al. Risk factors for progression of low-grade dysplasia in patients with Barrett's esophagus. Gastroenterology 2011.
10. di Pietro M, Bertani H, O'Donovan M, et al. Development and validation of confocal endomicroscopy diagnostic criteria for low-grade dysplasia in Barrett's esophagus. Clin Transl Gastroenterol 2019;10:e00014.
11. Gaddam S, Mathur SC, Singh M, et al. Novel probe-based confocal laser endomicroscopy criteria and interobserver agreement for the detection of dysplasia in Barrett's esophagus. Am J Gastroenterol 2011.
12. Sharma P, Meining AR, Coron E, et al. Real-time increased detection of neoplastic tissue in Barrett's esophagus with probe-based confocal laser endomicroscopy: final results of an international multicenter, prospective, randomized, controlled trial. Gastrointest Endosc 2011.
13. Canto MI, Anandasabapathy S, Brugge W, et al. In vivo endomicroscopy improves detection of Barrett's esophagus-related neoplasia: a multicenter international randomized controlled trial (with video). Gastrointest Endosc 2014.
14. Curvers WL, Van Vilsteren FG, Baak LC, et al. Endoscopic trimodal imaging versus standard video endoscopy for detection of early Barrett's neoplasia: a multicenter, randomized, crossover study in general practice. Gastrointest Endosc 2011.
15. Pietro MD, Boerwinkel DF, Shariff MK, et al. The combination of autofluorescence endoscopy and molecular biomarkers is a novel diagnostic tool for dysplasia in Barrett's oesophagus. Gut 2015.
16. Di Pietro M, Bird-Lieberman EL, Liu X, et al. Autofluorescence-directed confocal endomicroscopy in combination with a three-biomarker panel can inform management decisions in Barrett's esophagus. Am J Gastroenterol 2015;110:1549–1558.
17. Sharma P, Dent J, Armstrong D, et al. The development and validation of an endoscopic grading system for Barrett's esophagus: the Prague C & M criteria. Gastroenterology 2006;131:1392–1399.
18. Dixon MF. Gastrointestinal epithelial neoplasia: Vienna revisited. Gut 2002.
19. di Pietro M, Boerwinkel DF, Shariff MK, et al. The combination of autofluorescence endoscopy and molecular biomarkers is a novel diagnostic tool for dysplasia in Barrett's oesophagus. Gut 2015;64:49–56.

20. Boerwinkel DF, Shariff MK, Di Pietro M, et al. Fluorescence imaging for the detection of early neoplasia in Barrett's esophagus: old looks or new vision? Eur J Gastroenterol Hepatol 2014.
21. Pilonis ND, Killcoyne S, Tan WK, et al. Use of a Cytosponge biomarker panel to prioritise endoscopic Barrett's oesophagus surveillance: a cross-sectional study followed by a real-world prospective pilot. Lancet Oncol 2022.
22. Hadjinicolaou AV, van Munster SN, Achilleos A, et al. Aneuploidy in targeted endoscopic biopsies outperforms other tissue biomarkers in the prediction of histologic progression of Barrett's oesophagus: a multi-centre prospective cohort study. EBioMedicine 2020;56:102765.
23. Redston M, Noffsinger A, Kim A, et al. Abnormal TP53 predicts risk of progression in patients with Barrett's esophagus regardless of a diagnosis of dysplasia. Gastroenterology 2021.

Reprint requests
Address requests for reprints to: Massimiliano di Pietro, MD, MRC Cancer Unit, University of Cambridge, Cambridge Biomedical Campus, Box 197, CB20XZ, Cambridge, United Kingdom. e-mail: md460@cam.ac.uk; fax: (44) 1223763241.

Acknowledgments
The authors thank Bincy Alias, Irene Debiram-Beecham, and Tara Nuckcheddy (MRC Cancer Unit, University of Cambridge) for their help with patient recruitment; Nuria Galeano-Dalmau and Danial Hayward (MRC Cancer Unit, University of Cambridge) for assistance with immunohistochemistry staining; and Professor Rebecca Fitzgerald (MRC Cancer Unit, University of Cambridge) for critical comments on the manuscript and providing facilities and funding infrastructure for the molecular analyses.

CRediT Authorship Contributions
Mathew Vithayathil, MBBS MA(Hons)Cantab MRCP (Data curation: Lead; Formal analysis: Lead; Investigation: Supporting; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Supporting)
Ines Modolell, MD; PhD (Investigation: Equal; Writing – review & editing: Equal)
Jacobo Ortiz Fernández-Sordo, MD (Investigation: Equal; Writing – review & editing: Equal)
Dahmane Oukrif, MSc (Investigation: Equal; Writing – review & editing: Equal)
Apostolos Pappas, MD (Investigation: Equal; Writing – review & editing: Equal)
Wladyslaw Januszewicz, MD (Investigation: Equal; Writing – review & editing: Equal)
Maria O'Donovan, MD (Investigation: Equal)
Andreas Hadjinicolaou, MD, PhD (Investigation: Equal; Writing – review & editing: Equal)
Michele Bianchi, BSc (Methodology: Equal; Project administration: Equal; Writing – review & editing: Equal)
Adrienn Blasko, BSc (Investigation: Equal; Writing – review & editing: Equal)
Jonathan White, MBBS (Investigation: Equal; Writing – review & editing: Equal)
Philip Kaye, MBChB (Investigation: Equal; Writing – review & editing: Equal)
Marco Novelli, MBChB MSc PhD (Investigation: Equal; Writing – review & editing: Equal)
Lorenz Wernisch, PhD (Formal analysis: Equal; Writing – review & editing: Equal)
Krish Ragunath, MBBS, MD, MPhil (Conceptualization: Equal; Methodology: Equal; Supervision: Equal; Validation: Equal; Writing – review & editing: Equal)
Massimiliano di Pietro, MD (Conceptualization: Equal; Formal analysis: Equal; Funding acquisition: Lead; Investigation: Equal; Methodology: Equal; Resources: Lead; Supervision: Equal; Validation: Equal; Writing – original draft: Equal)

Conflicts of interest
This author discloses the following: Krish Ragunath is a consultant for Olympus. The remaining authors disclose no conflicts.

Funding
The study was funded by a Cancer Research UK Pump Priming research grant (M.d.P.). Additional funds were generously donated by the Cambridge Cancer Research Funds charity. Infrastructure support was received by the Experimental Cancer Medicine Centre and NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The laboratory of Professor Rebecca Fitzgerald is funded by a core program grant from the Medical Research Council (RG84369). The funding bodies had no role in the design or conduct of the research. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Transparency Statement
All anonymized individual participant data collected in the study will be available with publication. Study protocol and informed consent forms also will be made available. This will be available to researchers after submission of an approved proposal. All proposals should be directed to md460@cam.ac.uk, and requestors will need to sign a data access agreement.

Supplementary Methods

Patient-Reported Experience and Outcome Measures

Patient-reported experience using validated questionnaires was measured at baseline, and after each endoscopy. Distress and anxiety were measured using a 6-item state-trait anxiety inventory, previously used in endoscopy studies.[1] For the 6 state-traits (calm, tense, upset, relaxed, content, and worried), patients were assigned a rating as follows: not at all, somewhat, moderately, or very much. The overall procedure experience was assessed using a 10-point visual analogue scale (0 = worse, 10 = best). After completion of the second endoscopy, patients' preference between each arm was recorded. The pCLE diagnosis (dysplasia vs no dysplasia) was communicated to patients, immediately after the experimental endoscopy or once patients recovered from the sedation. For patients receiving sedation, the questionnaire was filled out at home and sent back by regular mail. Patient-reported experiences between experimental endoscopy and standard endoscopy were compared. The Q30 Wilcoxon signed rank sum test was used to compare STAI scores between experimental and standard arms. Visual analogue scores were compared using a paired *t* test.

Optical Probe-Based Confocal Laser Endomicroscopy Diagnosis

Before participating in the trial, 5 endoscopists underwent pCLE online training modules (http://www.cellvizio.net) until achieving at least 90% correct scoring in 10 consecutive video sets and then performed 5 pCLE procedures supervised by one of the expert pCLE endoscopists (M.d.P. or K.R.) at each institution. For a diagnosis of optical dysplasia by pCLE, 2 validated criteria sets were used[2,3] (Supplementary Methods). A diagnosis of HGD was made based on the presence of at least 2 of the following criteria[3]: saw-toothed epithelial surface, enlarged cells, pleiomorphic cells, nonequidistant glands, glands unequal in size and shape, and goblet cells not easily identified. A diagnosis of LGD required 3 of the following 6 criteria: dark nonround glands, irregular gland shape, lack of goblet cells, variable degree of darkness with sharp cut-off, value, variable size of cells, and cellular stratification.[2]

Molecular Biomarker Assays

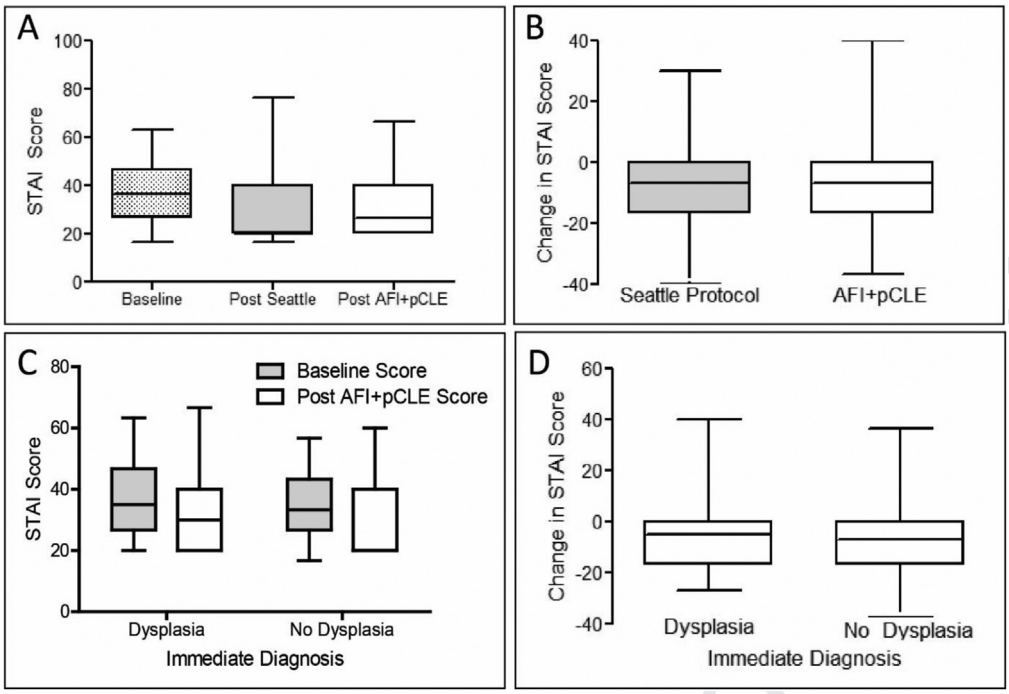
Immunohistochemistry was used to assess cyclin A Q31 (1:40; Novocastra) and p53 (p53 clone DO7, 1:50; Dakocytomation) expression with automated staining (BOND System; Leica Microsystems, Milton Keynes, UK). Cyclin A was scored by 2 of 5 independent investigators (M.d.P., A.B., A.P., M.V., and A.H.) and reviewed by a third investigator in cases of disagreement. Positive staining was considered a percentage of positive surface cells of 1% or greater.[4] p53 expression was scored by 2 investigators (M.O.'D. and P.K.); staining was reported as positive in case of strong focal staining or complete loss of staining, compared with the background expression.[5] Aneuploidy was assessed by image cytometry on cells isolated from frozen biopsy specimens.[6] The cell-cycle histogram was analyzed using ModFIT LT (Verity Software House, Topsham, ME).

Supplementary Results

We also looked at whether communication of optical diagnosis immediately after the procedure affected patient experience and anxiety levels. A significant decrease in anxiety scores was seen in postexperimental and poststandard arms compared with baseline (mean baseline score of 36.8 vs postexperimental score of 30.2; $P < .001$; vs poststandard arm score of 28.7; $P < .001$) (Supplementary Figure 1). There was no significant difference between the reduction in anxiety scores from baseline between the experimental and standard arms (-6.2 vs -8.2; $P = .33$). Furthermore, being immediately informed of a dysplasia diagnosis compared with waiting for histology results did not change patient anxiety scores, regardless of dysplasia status. Patients given a positive diagnosis of dysplasia did not show a significant change in anxiety from baseline compared with receiving a negative diagnosis. There was no difference in visual analogue scores between the experimental and standard arms (7.8 vs 7.8; $P = .98$).

Supplementary References

1. Williams J, Russell T, Durai D, et al. Effectiveness of nurse delivered endoscopy: findings from randomised multi-institution nurse endoscopy trial (MINuET). BMJ 2009. Q32
2. di Pietro M, Bertani H, O'Donovan M, et al. Development and validation of confocal endomicroscopy diagnostic criteria for low-grade dysplasia in Barrett's esophagus. Clin Transl Gastroenterol 2019;10:e00014.
3. Gaddam S, Mathur SC, Singh M, et al. Novel probe-based confocal laser endomicroscopy criteria and interobserver agreement for the detection of dysplasia in Barrett's esophagus. Am J Gastroenterol 2011.
4. Lao-Sirieix P, Lovat L, Fitzgerald RC. Cyclin A immunocytology as a risk stratification tool for Barrett's esophagus surveillance. Clin Cancer Res 2007.
5. Depledge DP, Evans KJ, Ivens AC, et al. Comparative expression profiling of Leishmania: modulation in gene expression between species and in different host genetic backgrounds. PLoS Negl Trop Dis 2009.
6. Dunn JM, Mackenzie GD, Oukrif D, et al. Image cytometry accurately detects DNA ploidy abnormalities and predicts late relapse to high-grade dysplasia and adenocarcinoma in Barrett's oesophagus following photodynamic therapy. Br J Cancer 2010;102:1608–1617.
7. di Pietro M, Boerwinkel DF, Shariff MK, et al. The combination of autofluorescence endoscopy and molecular biomarkers is a novel diagnostic tool for dysplasia in Barrett's oesophagus. Gut 2015;64:49–56.



Supplementary Figure 1. (*A*) STAI scores for patients at baseline and after Seattle and autofluorescence imaging (AFI)-guided probe-based confocal laser endomicroscopy (pCLE). (*B*) Change from baseline in STAI scores after Seattle and AFI + pCLE. (*C*) Baseline and post-AFI + pCLE STAI scores when an immediate diagnosis of dysplasia and no dysplasia was given after experimental endoscopy. (*D*) Changes in STAI scores after AFI + pCLE endoscopy from baseline for an immediate diagnosis of dysplasia compared with no dysplasia.

Supplementary Table 1. Per-Lesion Analysis of Diagnostic Accuracy of HRWLE, AFI, and AFI-Targeted Optical pCLE Diagnosis in the Experimental Arm

	HRWLE	AFI	AFI + pCLE	*P* value[a]
Dysplasia (all grades) (n = 27)				
Sensitivity	33.3	88.9	66.7	.01
Specificity	82.9	19.5	75.3	.02
High-grade dysplasia (n = 13)				
Sensitivity	38.5	92.3	69.2	.046
Specificity	82.3	19.3	73.2	.01

AFI, autofluorescence imaging; HRWLE, high-resolution white-light endoscopy; pCLE, probe-based confocal laser endomicroscopy.
[a]*P* value calculated for McNamar test for high-resolution endoscopy vs pCLE optical diagnosis.

Supplementary Table 2. By Operator: Per-Patient Analysis of Diagnostic Accuracy of Seattle Protocol Histology, AFI, AFI-Targeted Histology, and AFI-Targeted Optical pCLE Diagnosis

		All grades dysplasia		High-grade dysplasia	
	N	Sensitivity	Specificity	Sensitivity	Specificity
Operator 1	47	91.7	60.0	100	52.4
Operator 2	42	58.3	66.7	50.0	61.1
Operator 3	27	100.0	73.7	100	66.7
Operator 4	17	50.0	72.7	66.7	71.4
Operator 5	1	–	100	–	100

AFI, autofluorescence imaging; pCLE, probe-based confocal laser endomicroscopy.

Supplementary Table 3. Per-Patient Multivariate Logistic Regression Model for Predicting Histologic Dysplasia With AFI-Targeted pCLE, p53, Cyclin A, and Aneuploidy

	OR	95% CI	*P* value
pCLE	6.9	2.3–20.6	<.001
p53	13.1	3.6–47.5	<.001
Cyclin A	1.2	0.5–3.4	.67
Aneuploidy	2.6	0.8–8.6	.12

AFI, autofluorescence imaging; OR, odds ratio; pCLE, probe-based confocal laser endomicroscopy.