



Published in final edited form as:

Gastrointest Endosc. 2010 July ; 72(1): 19–24. doi:10.1016/j.gie.2010.01.053.

Preliminary accuracy and interobserver agreement for the detection of intraepithelial neoplasia in Barrett's esophagus with probe-based confocal laser endomicroscopy

Michael B. Wallace, MD, MPH, Prateek Sharma, MD, Charles Lightdale, MD, Herbert Wolfsen, MD, Emmanuel Coron, MD, Anna Buchner, MD, PhD, Monther Bajbouj, MD, Ajay Bansal, MD, Amit Rastogi, MD, Julian Abrams, MD, Julia E. Crook, MD, and Alexander Meining, MD, PhD

Jacksonville, Florida; Kansas City, Missouri; New York, New York, USA; Nantes, France; Munich, Germany

Abstract

Background—Confocal laser endomicroscopy (CLE) is a rapidly emerging method for in vivo imaging of the GI tract.

Objective—To determine the preliminary evaluation accuracy and interobserver agreement of probe-based CLE (pCLE) in Barrett's esophagus (BE).

Design—Prospective, double-blind review of pCLE images of 40 sites of BE tissue by using matching biopsies as the reference standard. A training set of 20 images with known histology was first reviewed to standardize image interpretation, followed by blinded review of 20 unknown images.

Setting—Eleven experts in BE imaging from 4 different endoscopy centers from the United States and Europe evaluated the images.

Patients—This study involved nonconsecutive patients undergoing BE surveillance or evaluation of high-grade intraepithelial neoplasia or early adenocarcinoma.

Intervention—Intravenous fluorescein pCLE imaging of each site within the BE segment, followed by matching biopsy.

Main Outcome Measurements—Sensitivity, specificity, and agreement for the pCLE diagnosis of high-grade intraepithelial neoplasia or carcinoma.

Results—In the validation set ($n = 20$), 11 cases had high-grade intraepithelial neoplasia or invasive carcinoma. The sensitivity for the diagnosis of neoplasia for the 11 endoscopists was 88% (range 6 of 11 to 11 of 11), and the specificity was 96% (range 7 of 9 to 9 of 9). There was substantial agreement on the pCLE diagnosis (86%, kappa 0.72; 95% confidence interval, 0.58–

Copyright © 2010 by the American Society for Gastrointestinal Endoscopy

Reprint requests: Michael B Wallace, MD, MPH, Professor of Medicine, Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL 32224. Current affiliations: Department of Gastroenterology and Hepatology, Mayo Clinic (M.B.W., H.W., A. Buchner, J.E.C.), Jacksonville, Florida, Veterans Administration Medical Center and Department of Gastroenterology and Hepatology, University of Kansas School of Medicine (P.S., A. Bansal, A.R.), Kansas City, Missouri, Department of Gastroenterology and Hepatology, Columbia University (C.L., J.A.), New York, New York, Department of Gastroenterology and Hepatology, Centre Hospitalier Universitaire Nantes (E.C.), Nantes, France, Department of Gastroenterology and Hepatology, Klinikum rechts der Isar, Technical University of Munich (M.B., A.M.), Munich, Germany.

DISCLOSURE: M. Wallace disclosed educational support below the level of federal reporting from Mauna Kea Technologies, P. Sharma disclosed an unspecified relationship with Mauna Kea Technologies, and C. Light-dale disclosed a consultant relationship with Olympus, America. All other authors disclosed no financial relationships relevant to this publication.

0.86). Endoscopists with prior pCLE experience had an overall sensitivity of 91% (all 10 of 11), specificity of 100% (all 9 of 9), and almost perfect agreement (92%, kappa 0.83; 95% confidence interval, 0.64–1.0).

Limitations—Small sample size and use of offline video sequences.

Conclusion—Results suggest that pCLE for the diagnosis of neoplasia in BE has very high accuracy and reliability.

Barrett's esophagus (BE) is the premalignant lesion for adenocarcinoma of the esophagus and esophagogastric junction. In this condition, the columnar mucosa replaces the squamous mucosa of the distal esophagus. The incidence of esophageal adenocarcinoma has been rapidly rising, increasing from 3-fold to 6-fold since 1990.¹ Patients with chronic GERD symptoms are advised to have endoscopic screening for the detection of BE. Random biopsies are also recommended if visible columnar epithelium appears in the distal esophagus. Once these random biopsies detect BE, patients are subsequently enrolled in a surveillance program. The presence of intraepithelial neoplasia (IEN) or early adenocarcinoma within a segment of BE is patchy; thus, standard endoscopy and random biopsies may fail to detect these lesions. New endoscopic imaging techniques to improve the accuracy of endoscopic diagnosis have been developed recently, and most are currently under evaluation.^{2,3}

The current method of random biopsy has several major limitations, including very low prevalence of IEN on a per-biopsy basis, leading to high cost per dysplasia detected; sampling errors; and poor interobserver agreement for dysplasia, even among expert pathologists. Importantly, the need for histology confirmation of neoplasia eliminates the ability to direct therapy during the index endoscopy because the endoscopists cannot "see" the location of disease. Thus, repeated endoscopies are needed, one for diagnosis, one for therapy.

Confocal laser endomicroscopy (CLE) is a novel imaging method that allows real-time micron-scale imaging of the GI tract, providing images very similar to those of traditional histology. CLE is available as probe-based confocal laser endomicroscopy (pCLE) or is integrated into a dedicated endoscope-based system. CLE has the theoretical potential to overcome many of the limitations of histology. In this study involving 11 experts with extensive experience with advanced endoscopic imaging of BE, we obtained a preliminary assessment of the accuracy and interobserver agreement of pCLE for the detection of neoplasia in BE.

METHODS

Assessment of accuracy and agreement

This study was a preliminary evaluation of pCLE in BE to establish baseline accuracy, image interpretation standards, and interobserver agreement. Eleven investigators were chosen, based on either prior experience with pCLE in BE (4 investigators) or extensive experience with other advanced endoscopic imaging of BE (7 investigators). Each investigator was provided with 40 images that were divided into a teaching set (10 with known IEN and 10 known to be without IEN). This was followed by a validation set of 20 unknown images. The exact prevalence of dysplasia in the validation set was not disclosed to the observers. The study was approved at the primary center (Klinikum Rechts der Isar), and each site obtained separate internal review board approval. All patients signed informed consent to participate in the research study.

pCLE image acquisition

We enrolled nonconsecutive patients undergoing BE surveillance and patients referred for evaluation and treatment of BE-associated, high-grade IEN or early adenocarcinoma. Standard video endoscopies with both high-definition white-light and narrow-band imaging (Olympus Exera GIF H 180; Olympus, Center Valley, Penn) were performed first. pCLE images were obtained from both suspicious and nonsuspicious areas. In order to precisely match each imaging site to the subsequent biopsy site, an essential element for any research study, each site of imaging was marked with a cautery “dot” using argon plasma coagulation, immediately before pCLE imaging. Intravenous fluorescein was injected at a dose of 0.1 mL/kg of 1% (Alcon Pharmaceutical, Hünenberg, Switzerland). Imaging was performed within the first 1 to 8 minutes after fluorescein injection, as per prior dose-timing studies.⁴ All images were obtained with the ultra High Definition pCLE probe (UHD; Gastroflex, Mauna Kea Technologies, Paris, France). The site was then sampled with a standard biopsy forceps, and all histology was reviewed by an expert GI pathologist, by using the updated Vienna classification of intraepithelial neoplasia.⁵

Images were selected based on the presence of interpretable epithelial images as determined by the primary investigator (A.M.). Each video consisted of approximately 20 to 30 seconds of pCLE imaging obtained from the BE segment.

Videos were obtained in a proprietary .mkt format and converted to a universal (.avi) format, with no compression for distribution to each observer. All observers viewed and scored the validation set videos without knowledge of the histology or the interpretation of the other investigators.

The pCLE criteria for neoplasia were adapted from Pohl et al.⁶ Overall, neoplasia was classified when at least 2 of the 5 vascular or cellular criteria were present: irregular epithelial lining, variable width of the epithelial lining, fusion of glands, presence of dark areas (decreased uptake of fluorescein), or irregular vascular pattern.

Accuracy was assessed by using pooled estimates of sensitivity and specificity across the 11 observers. Interobserver agreement was summarized as a percentage (average agreement across all pairs of observers). A kappa statistic was calculated to assess the degree of agreement that this represented, beyond that expected by chance alone. A level of agreement consistent with chance would have a kappa estimate around zero; perfect agreement would yield a kappa estimate of 1.0. Landis and Koch⁷ described their “arbitrary benchmark interpretations” as follow: scores of 0.41 to 0.60 moderate, 0.61 to 0.80 substantial, and 0.81 to 1.00 almost perfect. It is important to recognize that these descriptors are imprecise and represent only qualitative methods to describe numerical data.

RESULTS

Forty sites with BE from 5 separate patients (3 male, 2 female; ages 38–86 years) were imaged with pCLE and had matching biopsies. No complications occurred during the pCLE procedures. In the training set, 10 sites had high-grade IEN or invasive cancer, and 10 had Barrett’s metaplasia with no IEN.

In the validation set, 11 sites had high-grade IEN, and 9 had no IEN. Representative still images of BE without IEN, high-grade IEN, and invasive carcinoma are shown in Figures 1–4, with links to video. Using the observers’ overall assessments, the 11 experts had very high sensitivity and specificity for the diagnosis of neoplasia in these 20 sites from 5 patients (Table 1). The 4 observers with prior pCLE experience (experts) had higher accuracy, although the small, pilot nature of the study did not allow for any formal statistical

comparison among observers. The overall interobserver agreement was substantial, estimated at 86% with a kappa estimate of 0.72 (95% confidence interval, 0.58–0.86), and almost perfect for the 4 most experienced pCLE observers (Table 1).

In addition to the overall assessment of neoplasia versus no neoplasia, the observers also rated individual features suggestive of neoplasia, such as irregular epithelial thickness, epithelial inhomogeneity, dark epithelial structures (lack of fluorescein uptake), crypt/villi fusion, and irregular vessels. These individual features had good specificity but lower sensitivity, and none of them appeared to compete with the overall diagnostic assessment (Table 2).

DISCUSSION

This preliminary study suggests that when pCLE is used by a broad group of endoscopists with specialized interest in advanced imaging, it is an accurate and reliable method for the diagnosis of IEN in BE, and experts (prior experience with pCLE) perform better than nonexperts. The overall diagnosis, based on the presence of at least two of the individual vascular and epithelial features, appears to be more accurate than any individual diagnostic feature.

Multiple advanced optical methods for guiding biopsy in BE have been evaluated in the past. These can be divided into broad-field methods that can guide biopsy and small-field methods, including pCLE, that can confirm or exclude neoplasia in suspicious sites. Small-field methods such as pCLE are not well-suited to surveying large areas of tissue such as long segments of BE and should ideally be combined with broad-field (“red flag”) methods. Traditionally, staining agents such as indigo carmine have been used to enhance the contrast of the mucosal surface during magnifying endoscopy. Narrow-band imaging (NBI) is a novel technique that increases the mucosal contrast without the use of dyes. In a recent randomized, crossover study, NBI was shown to be at least as effective as chromoendoscopy for the detection of neoplasia in BE.⁸ Because NBI is a simple technique that uses only modified optical filters, it may eliminate the need for staining agents. Other studies also have evaluated NBI for the detection of high-grade dysplasia and cancer and have found NBI to be beneficial for the detection of these lesions. This technique suffers from a relatively low specificity, but it can be a good screening examination to evaluate the entire surface area of the BE.^{9–11} In our recent, randomized, double-blind, tandem trial of standard white-light endoscopy versus high-definition NBI, we found significantly more dysplasia and higher grades of dysplasia, and fewer biopsies were required with NBI compared with white-light endoscopy. The main limitation of NBI (and other methods such as autofluorescence) is the relatively low specificity, such that most “positive” sites by NBI are actually false positives.¹²

These data suggest that NBI potentially can be used as a broad-field evaluation technique (red-flag technique) with improvement over white-light endoscopy, but the sensitivity and specificity are still insufficient to replace random biopsy or to avoid biopsy confirmation of all abnormal sites.

pCLE is a new method that provides cross-sectional diagnostic microscopic views of the mucosa with the help of a fibered laser probe. It is possible to image individual cells and subcellular structures, thus a statement concerning the tissue type and neoplasia diagnosis seems to be possible. Meining et al¹³ conducted a single-center study on the feasibility of the pCLE method to detect malignant and premalignant modifications in the GI tract of 47 patients (34 with neoplasia). The investigators estimated that this method has 92% accuracy in the detection of a neoplasia, as compared with conventional histopathology.¹³ The

disadvantage of this study was the fact that approximately one fifth of the pCLE sequences did not have a sufficient quality to enable interpretation of results. This disadvantage mostly can be explained by the design of the study. The investigators used probes with a confocal depth of only 10 μm (so-called, surface probes). As far as the fluorophore is concerned, cresyl violet was used topically after mucolysis with N-acetylcysteine. The disadvantage of this method is that some superficial mucus remains, and contact bleeding can degrade the quality of the picture.¹⁴ This problem can be solved by using probes with a deeper confocal plane and by using a dye injected intravenously, such as fluorescein. Preliminary data from our group, interpreted by using intravenous fluorescein and deeper imaging probes in colorectal neoplasia, confirmed significantly better image quality and reproducibility with this method.¹⁵ In addition, the single-center, single-endoscopist data, although encouraging, warrants validation among a broader group of users.

Likewise, after the injection of fluorescein, Kiesslich et al¹⁶ examined 63 patients with BE (15 with neoplasia) with endoscope-based confocal laser endomicroscopy. In a single-center study, using this method and an extremely experienced endomicroscopist, the authors were able to achieve surprisingly high accuracy, both in the diagnosis of Barrett's metaplasia and Barrett's neoplasia (96.8% and 97.4%, respectively). A major limitation of this work is the required use of a dedicated endoscope-based confocal laser endomicroscopy endoscope (Pentax, Ft Wayne, NJ) and topical staining with acriflavine, a potential carcinogen. pCLE with fluorescein alone overcomes these limitations, allowing safe pCLE with any endoscopic system, with the ability to perform ad hoc pCLE whenever needed.

In a single-center, preliminary study, Pohl et al⁶ evaluated the pCLE system for the characterization of neoplastic areas in patients with BE. This study evaluated the preliminary accuracy of pCLE for high-grade dysplasia and early esophageal adenocarcinoma in patients with BE by using a lower-definition probe (GastroFlex type Z; Mauna Kea Technologies). They evaluated 296 biopsy sites from 38 patients. Carcinoma or high-grade dysplasia was detected in 6.4% of biopsies. The overall accuracy of pCLE for the two different observers was 88% and 93%, with sensitivity of 75% and 80%, specificity of 89% and 94%, positive predictive value of, at best, 44.4%, and negative predictive value of, at best, 98.8%. There was a reasonable degree of interobserver agreement between the two endoscopists (kappa 0.6).

The main strength of the present study relative to other published data is the inclusion of a large group of endoscopists. Although the majority of these endoscopists can be considered experts in BE imaging, as evidenced by a track record of published clinical trials, only 4 had experience with pCLE before this study. The high performance, even of those not considered experts, demonstrated that rapid training in image interpretation is feasible. It is likely that these advanced imaging methods will first be applied in referral centers, like those represented in this study; thus, our results represent the typical setting where pCLE may be used in practice.

One of the limitations of the study is that we evaluated image interpretation accuracy by using offline video sequences only. Real-time image interpretation by multiple endoscopists, although desirable, is not technically feasible in most settings. Use of recorded representative videos is an accepted method of interpreting endoscopic accuracy and reliability.¹⁷ There is the possibility that real-time information could affect accuracy in either direction. Availability of other clinical information such as a history of neoplasia, or the endoscopic images themselves, could potentially improve diagnostic accuracy; however, the pressure to interpret images in real time "on the fly" while doing the endoscopy may be more difficult than in the more controlled offline setting. The same group of endoscopists is carefully studying both of these issues in a large, prospective, multicenter trial. Additionally,

the preferential inclusion of high-quality video clips may limit the applicability of the findings of this preliminary study. The use of only nondysplastic and high-grade dysplasia samples also has the potential to overestimate the level of agreement and accuracy. In this initial study, we evaluated the technology under ideal circumstances; however, further study is clearly needed to see whether these results are reliable in a more typical clinical setting with varying degrees of dysplasia. Last, the small sample size of the validation set with denominators (11 and 9, respectively) for the estimation of sensitivity and specificity mean that these quantities were estimated with very low precision; therefore, these initial suggestive findings require confirmation with larger studies.

Overall, the data from our study suggest that pCLE has the potential to accurately and reliably diagnose IEN in patients with BE. The degree of interobserver agreement was high among all endoscopists, even among those with no previous experience with pCLE, suggesting a short learning curve in the interpretation of images. This will positively impact the application of confocal endomicroscopy as a useful tool in clinical practice.

Abbreviations

BE	Barrett's esophagus
CLE	confocal laser endomicroscopy
IEN	intraepithelial neoplasia
NBI	narrow-band imaging
pCLE	probe-based confocal laser endomicroscopy

References

1. Cameron AJ. Epidemiology of columnar-lined esophagus and adenocarcinoma. *Gastroenterol Clin North Am.* 1997; 26:487–94. [PubMed: 9309399]
2. Wang KK, Sampliner RE. Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *Am J Gastroenterol.* 2008; 103:788–97. [PubMed: 18341497]
3. Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett's esophagus. *Am J Gastroenterol.* 2002; 97:1888–95. [PubMed: 12190150]
4. Becker V, von Delius S, Bajbouj M, et al. Intravenous application of fluorescein for confocal laser scanning microscopy: evaluation of contrast dynamics and image quality with increasing injection-to-imaging time. *Gastrointest Endosc.* 2008; 68:319–23. [PubMed: 18436217]
5. Schlemper RJ, Riddell RH, Kato Y, et al. The Vienna classification of gastrointestinal epithelial neoplasia. *Gut.* 2000; 47:251–5. [PubMed: 10896917]
6. Pohl H, Rösch T, Vieth M, et al. Miniprobe confocal laser microscopy for the detection of invisible neoplasia in patients with Barrett's oesophagus. *Gut.* 2008; 57:1648–53. [PubMed: 18755886]
7. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977; 33:159–74. [PubMed: 843571]
8. Curvers W, Baak L, Kiesslich R, et al. Chromoendoscopy and narrow-band imaging compared with high-resolution magnification endoscopy in Barrett's esophagus. *Gastroenterology.* 2008; 134:670–9. [PubMed: 18242603]
9. Kara MA, Ennahachi M, Fockens P, et al. Detection and classification of the mucosal and vascular patterns (mucosal morphology) in Barrett's esophagus by using narrow band imaging. *Gastrointest Endosc.* 2006; 64:155–66. [PubMed: 16860062]
10. Kara MA, Peters FP, Fockens P, et al. Endoscopic video-autofluorescence imaging followed by narrow band imaging for detecting early neoplasia in Barrett's esophagus. *Gastrointest Endosc.* 2006; 64:176–85. [PubMed: 16860064]

11. Kara MA, Peters FP, Rosmolen WD, et al. High-resolution endoscopy plus chromoendoscopy or narrow-band imaging in Barrett's esophagus: a prospective randomized crossover study. *Endoscopy*. 2005; 37:929–36. [PubMed: 16189764]
12. Wolfsen HC, Crook JE, Krishna M, et al. Prospective, controlled tandem endoscopy study of narrow band imaging for dysplasia detection in Barrett's esophagus. *Gastroenterology*. 2008; 135:24–31. [PubMed: 18442484]
13. Meining A, Saur D, Bajbouj M, et al. In vivo histopathology for detection of gastrointestinal neoplasia with a portable, confocal miniprobe: an examiner blinded analysis. *Clin Gastroenterol Hepatol*. 2007; 5:1261–7. [PubMed: 17689297]
14. Becker V, Vercauteren T, von Weyhern CH, et al. High-resolution miniprobe-based confocal microscopy in combination with video mosaicing (with video). *Gastrointest Endosc*. 2007; 66:1001–7. [PubMed: 17767932]
15. Buchner A, Murli K, Wolfsen H, et al. High resolution confocal endomicroscopy probe system for in vivo diagnosis of colorectal neoplasia [abstract]. *Gastroenterology*. 2008; 135:295.
16. Kiesslich R, Gossner L, Goetz M, et al. In vivo histology of Barrett's esophagus and associated neoplasia by confocal laser endomicroscopy. *Clin Gastroenterol Hepatol*. 2006; 4:979–87. [PubMed: 16843068]
17. Wallace MB, Hawes RH, Durkalski V, et al. The reliability of EUS for the diagnosis of chronic pancreatitis: interobserver agreement among experienced endosonographers. *Gastrointest Endosc*. 2001; 53:294–9. [PubMed: 11231386]

Take-home Message

- Probe-based confocal laser endomicroscopy is a rapidly emerging imaging method that has the potential to replace random biopsy. Conditions with low yield for disease, such as surveillance in Barrett's esophagus, are good candidates for application of this method. This study suggests high levels of preliminary accuracy and interobserver agreement from a large number of experts in Barrett's esophagus imaging.

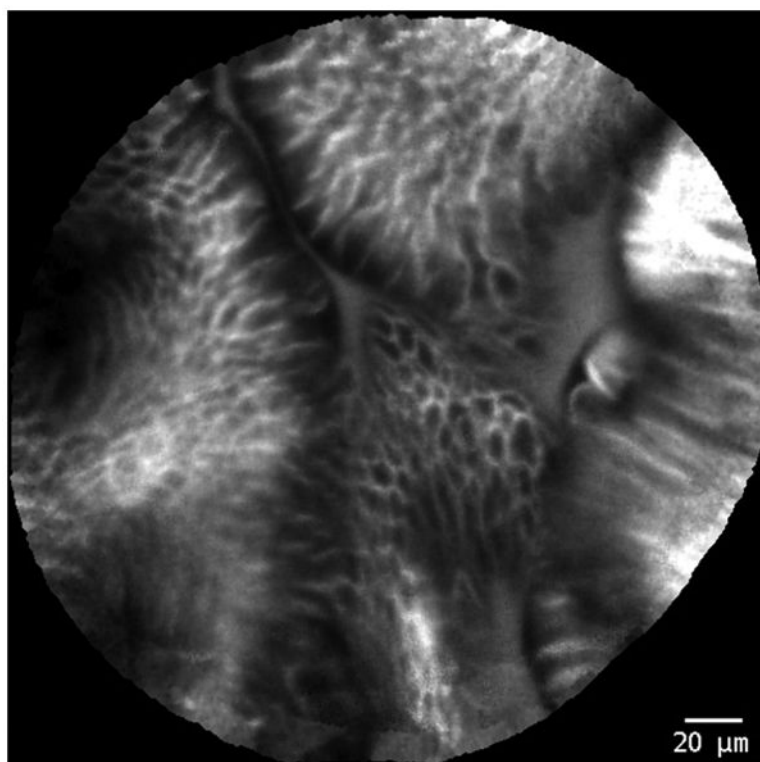


Figure 1. Probe-based confocal laser endomicroscopy images of Barrett's esophagus without neoplasia.

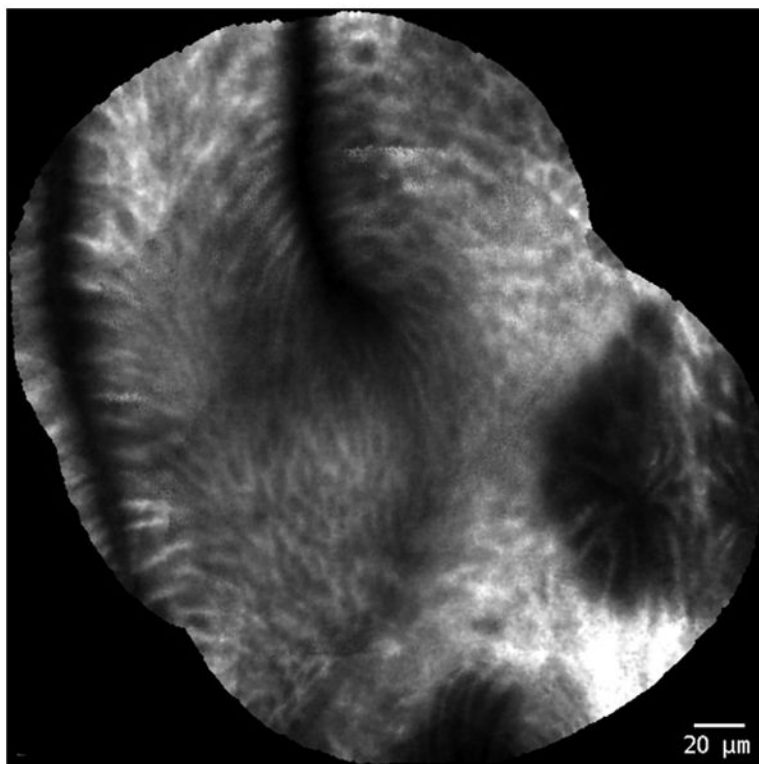


Figure 2. Probe-based confocal laser endomicroscopy images of Barrett's esophagus without neoplasia.

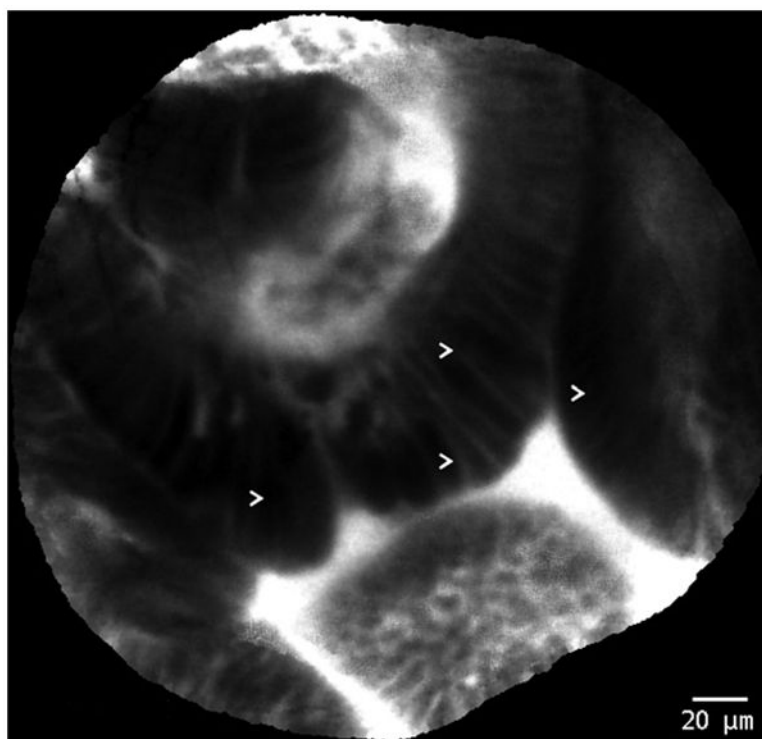


Figure 3. Probe-based confocal laser endomicroscopy images of Barrett's esophagus with high-grade intraepithelial neoplasia. Arrows (>) indicate areas of dark, irregularly thickened, epithelial borders that are characteristic of dysplasia.

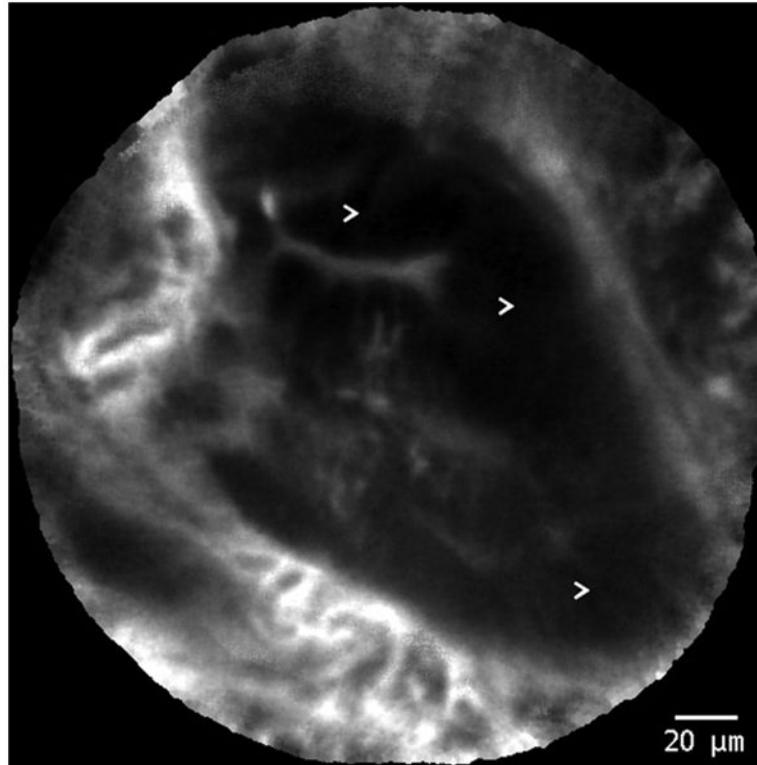


Figure 4. Probe-based confocal laser endomicroscopy images of Barrett's esophagus with high-grade intraepithelial neoplasia. Arrows (>) indicate areas of dark, irregularly thickened, epithelial borders that are characteristic of dysplasia.

TABLE 1

Accuracy of overall assessment of neoplasia by each of the 11 endoscopists (percentage and range of absolute values)

	pCLE experienced observers (n = 4)	pCLE inexperienced observers (n = 7)	10 of 11 observers
Sensitivity	91% (all 10/11)	87% (6/11–11/11)	88%
Specificity	100% (all 9/9)	94% (7/9–9/9)	96%
Accuracy	95% (all 19/20)	90% (14/20–20/20)	92%
Agreement	92%	82%	86%
Kappa (95% CI)	0.83 (0.64–1.00)	0.64 (0.48–0.80)	0.72 (0.58–0.86)

pCLE, probe-based confocal laser endomicroscopy; CI, confidence interval.

TABLE 2

Accuracy of individual features of neoplasia (percentage and range of absolute values)

Criterion	Sensitivity % (range)	Specificity % (range)	Accuracy % (range)
Irregular epithelial thickness	63 (3/11–10/11)	96 (7/9–9/9)	78 (12/20–19/20)
Epithelial inhomogeneity	75 (3/11–11/11)	93 (7/9–9/9)	83 (12/20–20/20)
Villi/crypt fusion	52 (0/11–10/11)	96 (7/9–9/9)	72 (8/20–19/20)
Dark epithelial boarder	63 (1/11–10/11)	94 (7/9–9/9)	77 (10/20–19/20)
Irregular vessels	50 (0/11–10/11)	97 (8/9–9/9)	71 (9/20–19/20)