Background

The current standard for surveillance of Barrett’s esophagus (BE) is to perform white-light endoscopy with targeted biopsies of any endoscopically visible lesions and random four-quadrant biopsies every 1–2 cm of the BE segment (i.e., Seattle protocol) [1]. Unfortunately, this strategy is labor intensive, may miss 10–50% of esophageal neoplasms, and may increase the risk of bleeding from several biopsies [2,3]. Additionally, multiple random biopsies may not be cost-effective, given the low absolute incidence of esophageal adenocarcinoma in patients with BE (0.4–0.5%) and the need for additional procedures for endoscopic eradication if dysplasia is detected [4,5].

Confocal laser endomicroscopy (CLE) is a novel endoscopic technique that permits real-time in vivo histologic assessment of esophageal mucosa during upper endoscopy. Prospective studies comparing the accuracy of CLE to 4-quadrant biopsies in routine clinical practice are lacking.

Methods

Consecutive patients with BE underwent high-definition white light and narrow-band imaging followed by CLE and targeted biopsy or mucosal resection. Four-quadrant biopsies were obtained during the same session. Baseline variables, real-time CLE interpretation, and histology results were prospectively recorded. Blinded expert review of CLE sequences and histology specimens was performed. A sample size of 64 patients was calculated a priori based on 3% estimated prevalence of high-grade dysplasia (HGD) or cancer.

Results

In total, 66 patients were included in the study. The prevalence of HGD or cancer was 4.55%. Both real-time and blinded CLE correctly identified all cases of cancer. For the primary outcome, real-time CLE was 98% specific but only 67% sensitive for HGD/cancer compared to non-blinded pathologist interpretation. For HGD and cancer, inter-observer agreement was substantial between real-time and blinded endomicroscopists (kappa = 0.6). CLE identified dysplasia in 75% of cases where both blinded and unblinded pathology interpretation was low-grade dysplasia.

Conclusions

CLE demonstrates high specificity for detecting dysplasia and cancer, but lower sensitivity may limit its utility in routine BE surveillance. CLE may have a role in confirming LGD in real-time before eradication therapy.
Technology is based on the principle of illuminating a tissue with a low-power laser and then detecting fluorescent light reflected from the tissue [6]. The laser light is focused at a selected depth and reflected light is then refocused on the detection system by the same lens [7]. The light reflected and scattered at other geometric angles from the illuminated tissue is excluded from detection, which dramatically increases the spatial resolution of CLE [7]. Commercially available CLE is based on tissue fluorescence, with a topical or intravenous contrast agent.

Several societies have endorsed pCLE in patients undergoing surveillance of BE [8, 9]. These recommendations are based on studies that compared pCLE to white-light endoscopy, used an endoscope-based version (eCLE) that is not commercially available, marked tentative biopsy sites with argon plasma coagulation (APC), or used pCLE along with autofluorescence imaging and genetic panel testing [10–13]. To our knowledge, prospective studies comparing pCLE to random biopsies in routine clinical practice (i.e., comparative effectiveness studies) are lacking.

To eliminate the need for random biopsies, the American Society of Gastrointestinal Endoscopy (ASGE) initiative for the Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) recommends that an imaging technology must demonstrate a per-patient sensitivity of at least 90% and specificity of at least 80% when compared to random biopsies [14]. The aim of this study was to prospectively assess if pCLE met the PIVI criteria for detecting dysplasia and cancer in routine clinical practice among patients undergoing endoscopic surveillance of BE.

**Methods**

**Patients and procedures**

The institutional review board at the Hunter Holmes McGuire Veterans Affairs Medical Center approved the study. Consecutive patients referred for surveillance endoscopy for BE underwent high-definition white-light endoscopy (HD-WLE) and narrow-band imaging (NBI) using an Olympus GIF-HQ190 adult gastroscope. Areas suspicious for dysplasia were identified on NBI based on the presence of irregular mucosal patterns, irregular vascular patterns, or abnormal blood vessels [15]. Following visual examination, pCLE examination was performed using a 2.5 mm gastroflex ultra-high-definition probe (Cellvizio G1 system, Mauna Kea, Paris, France) passed through the working channel of the gastroscope. The pCLE probe was placed gently on the mucosa after intravenous injection of 2.5 mL of 10% fluorescein. A transparent cap was fitted to the distal end of the scope to assist with probe stabilization. Video sequences were obtained from any areas felt to be suspicious on HD-WLE and NBI as well as in four quadrants at 1-cm intervals. Two gastroenterologists with at least 3 months of training in pCLE performed all procedures and interpreted pCLE recordings during the procedure (TS, PM). The investigators interpreted pCLE recordings based on the Miami classification as follows: non-dysplastic Barrett’s esophagus (uniform villiform architecture, columnar cells, dark goblet cells) (Fig. 1); adenocarcinoma (disorganized or absent villiform structures and crypts, dark columnar cells, dilated irregular vessels) (Fig. 2) [16]; dysplasia (villiform structures, dark irregularly thickened epithelial borders, dilated irregular vessels) (Fig. 3).

Endoscopists obtained targeted biopsies or performed endoscopic mucosal resection (EMR) of areas suspicious for
HGD or cancer on HD-WLE, NBI, or pCLE. Subsequently, they obtained random four-quadrant biopsies at 1–2 cm intervals as suggested by the major American gastroenterology societies [1, 17]. In patients with diminutive islands or tongues of suspected Barrett’s esophagus (i.e. ≤1 cm), the Seattle protocol was not feasible, so two to three random biopsies were obtained. Areas that were already sampled during targeted biopsy were not re-sampled while obtaining random biopsies. Baseline variables and real-time pCLE interpretation were prospectively recorded. As is routine at our institution, a single pathologist interpreted biopsies apart from when dysplasia or malignancy was suspected, in which case a second pathologist confirmed the diagnosis. Non-blinded pathology results were prospectively recorded.

**Blinded pCLE and pathology review**

A single expert endomicroscopist (AZ) who had performed > 1000 pCLE procedures for BE reviewed all pCLE video sequences. The expert endomicroscopist was blinded to endoscopic images, real-time pCLE interpretation, and pathology interpretations. Subsequently, two endomicroscopists (TS, PM) reviewed video sequences for patients identified as having dys-
plasia or cancer on pCLE to assess whether they met revised pCLE criteria developed and validated by Gaddam et al. [18]. These criteria are as follows: epithelial surface: saw-toothed; cells: enlarged; cells: pleomorphic; glands: not equidistant; glands: unequal in size and shape; goblet cells: not easily identified. An expert pathologist (RL) reviewed all histopathology specimens. The pathologist was blinded to endoscopy findings, blinded and non-blinded pCLE interpretation, and non-blinded histology interpretation. Standardized criteria were used for blinded pathology interpretation [19].

Sample size
The study was a retrospective assessment of a prospectively maintained endomicroscopy database. The primary hypothesis was that specificity of real-time pCLE at detecting HGD or cancer assessed on a per patient basis is not less than 80% when compared to random biopsy. Assuming a 3% prevalence of high grade dysplasia or cancer in our tertiary referral Veterans Affairs Medical Center, expected power of 80%, and a 0.05 level of significance, we estimated a sample size of 64 patients would be necessary to evaluate the specificity of real-time pCLE to unblinded pathology review [5].

Statistical analyses
Statistical analyses were performed using SPSS IBM version 24 (Chicago, IL, USA). Baseline variables were recorded as frequencies and percentages. Accuracy was assessed by calculating sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) of pCLE ± targeted biopsy compared to random biopsy. Cohen’s kappa statistic was calculated to assess the degree of agreement between real time and blinded endomicroscopists as well as between blinded and unblinded pathologists [20]. Conventionally, Kappa scores of 0.41 – 0.6 represent moderate agreement, 0.61 – 0.79 represent substantial agreement, and 0.8 – 1.00 represent nearly perfect agreement.

Results
A total of 66 patients were included in the study. Procedures were performed from December 2014 to October 2016. The median duration for pCLE examination was 7 minutes (▶ Table 1). The median age was 66 years, and the majority of patients were white men. Mean Barrett’s segment length was C2M3. Fifty-eight patients (88%) were on a proton pump inhibitor, 71% reported a history of cigarette smoking, and 86% were overweight or obese (▶ Table 1).

The overall prevalence of high grade dysplasia or cancer was 4.55% (2 cancers, 1 high grade dysplasia). Both patients with cancer had visible areas of mucosal irregularity on HD-WLE and NBI. For the primary outcome, accuracy of real-time pCLE for diagnosing HGD/cancer compared to non-blinded pathologist interpretation was as follows: sensitivity 67%, specificity 98%, negative predictive value 98%, and positive predictive value 67% (▶ Table 2). Both real-time and blinded pCLE correctly identified all cases of cancer. One patient with a flat diminutive tongue of salmon colored mucosa (COM1) interpreted as nondysplastic BE on real-time and blinded pCLE examination was found to have HGD on random biopsy. The diagnosis of HGD was confirmed on subsequent EMR, and on blinded pathology review. For HGD and cancer, inter-observer agreement was moderate between real-time and blinded endomicroscopists (kappa = 0.6), and was perfect between blinded and non-blinded pathologists (kappa = 1).

The prevalence of LGD in our cohort varied substantially from 6% to 29% depending on the modality (pCLE vs. biopsy) and physician (blinded vs. non-blinded) (▶ Table 3). Real-time pCLE identified LGD in 11 patients (17%), whereas blinded pCLE review diagnosed LGD in only 6% of patients. When two non-blinded pathologists evaluated the specimens, LGD was found in 8% of patients, whereas blinded histology review using standardized criteria identified dysplasia in 29% of patients. Specificity of real-time pCLE for LGD was greater than 80%.

### Table 1 Baseline variables (n = 66).

| Variable                                | Value          |
|-----------------------------------------|----------------|
| Median age, years                       | 66 (range 44 – 73) |
| Gender                                  | Male (98%)     |
| Race                                    | White (92%)    |
| Median BMI                              | 29 (range 17 – 46) |
| Mean length of Barrett’s esophagus, years | C2 (range C0 – C15) |
| Median duration of Barrett’s esophagus, minutes | 7 (range 2 – 26) |
| Median duration of pCLE exam, minutes   | 7 (range 2 – 26) |
| Proton pump inhibitor use (%)           | Yes 58 (88%)   |
| Current smoker (%)                      | Yes 21 (32%)   |
| Prior smoker (%)                        | Yes 47 (71%)   |
| Mean hiatal hernia size, cm             | 2 (range 0 – 9) |

BMI, body mass index; pCLE, probe-based confocal laser endomicroscopy.

### Table 2 Accuracy of probe-based confocal laser endomicroscopy (pCLE) compared to histology for high grade dysplasia or cancer.

|                   | Non-blinded pathologist interpretation |
|-------------------|----------------------------------------|
| Real-time pCLE interpretation | Sensitivity 67% (CI 99 – 99%) |
|                   | Specificity 98% (CI 91 – 100%) |
|                   | NPV 98% (CI 93 – 100%) |
|                   | PPV 67% (CI 20 – 95%) |
| Blinded pCLE interpretation     | Sensitivity 67% (CI 99 – 99%) |
|                   | Specificity 94% (85 – 98%) |
|                   | NPV 33% (12 – 63%) |
|                   | PPV 98% (92 – 100%) |

PPV, positive predictive value; NPV, negative predictive value; Agreement was 100% between non-blinded and blinded pathologists for HGD and cancer.
when assessed against both non-blinded and blinded pathology interpretations; specificity of blinded pCLE interpretation for LGD was greater than 90% (▶ Table 2). Sensitivity of pCLE for LGD was low when compared to random biopsies, particularly in the blinded pCLE group (▶ Table 3). For LGD, inter-observer agreement was poor between real-time and blinded endomicroscopists (kappa = 0.2) as well as between blinded and non-blinded pathologists (kappa = 0.2). Among the 20 patients who had LGD on either unblinded or blinded pathology review, only three patients had visible areas of nodularity or irregularity on HD-WLE or NBI (▶ Table 4). Real-time pCLE identified LGD in three of four patients (75%) who were diagnosed as LGD by both unblinded and blinded pathologists. All patients identified as having dysplasia or cancer on real-time or blinded pCLE review met at least one of the revised criteria proposed by the ASGE Technology Committee [7].

**Discussion**

Sharma et al. demonstrated improved sensitivity of pCLE compared to WLE [12]. However, in clinical practice, the utility of the technology would rest on its ability to eliminate the need for random biopsies, which are time consuming with resulting low adherence [21]. The aim of our study was to determine whether pCLE met PIVI criteria to consider replacing random biopsies for surveillance of BE in clinical practice [14]. The technology did meet the primary outcome of specificity >80% for HGD and cancer when compared to random biopsies. However, both patients with cancer had visible areas of mucosal irregu-

| Table 3 | Accuracy of probe-based confocal laser endomicroscopy (pCLE) compared to histology for low grade dysplasia. |
|-----------------|------------------------------------------------------------------------------------------------------------------|
| **Non-blinded pathologist interpretation** (n=5) | **Blinded pathologist interpretation** (n=19) |
| **Real-time pCLE interpretation** (n=9) | Sensitivity 60% Specificity 87% NPV 96% PPV 27% | Sensitivity 32% Specificity 89% NPV 76% PPV 55% |
| **Blinded pCLE interpretation** (n=4) | Sensitivity 0% Specificity 93% NPV 92% PPV 0% | Sensitivity 11% Specificity 96% NPV 73% PPV 50% |

PPV, positive predictive value; NPV, negative predictive value.

| Table 4 | pCLE and targeted biopsy/mucosal resection findings in patients with nodularity or irregularity on high definition white light or narrow-band imaging. |
|-----------------|-------------------------------------------------------------------------------------------------|
| **Patient** | **Sampling method** | **pCLE interpretation** | **Pathology interpretation** |
| 1 | Biopsy | Real-time Adenocarcinoma Blinded Adenocarcinoma | Unblinded Adenocarcinoma Blinded Adenocarcinoma |
| 2 | Mucosal resection | Real-time HGD Blinded HGD | Unblinded No Barrett’s esophagus Blinded No Barrett’s esophagus |
| 3 | Mucosal resection | Real-time LGD Blinded LGD | Unblinded No Barrett’s esophagus Blinded |
| 4 | Biopsy | Real-time NDB Blinded NDB | Unblinded NDB Blinded NDB |
| 5 | Biopsy | Real-time LGD Blinded HGD | Unblinded NDB Blinded NDB |
| 6 | Mucosal resection | Real-time LGD Blinded NDB | Unblinded Indefinite Blinded LGD |
| 7 | Mucosal resection | Real-time HGD Blinded LGD | Unblinded Indefinite Blinded LGD |
| 8 | Mucosal resection | Real-time NDB Blinded NDB | Unblinded No Barrett’s esophagus Blinded No Barrett’s esophagus |
| 9 | Biopsy | Real-time Adenocarcinoma Blinded HGD | Unblinded Adenocarcinoma Blinded Adenocarcinoma |

NDB, non-dysplastic Barrett’s esophagus; LGD, low grade dysplasia; HGD, high grade dysplasia.
larity or nodularity on HD-WLE or NBI, so pCLE did not provide any incremental benefit.

Additionally, pCLE did not meet the >90% sensitivity threshold for dysplasia and cancer recommended by the PIVI initiative. Sensitivity was not the primary outcome, so we cannot rule out the possibility that the study was underpowered to minimize type II error. Yet, our results are consistent with those of Bajbouj et al. [10], who marked tentative biopsy sites using argon plasma coagulation (APC), assessed the sites with pCLE, and then obtained biopsies from the sites. In their study, pCLE demonstrated high specificity but sensitivity of only 12–28%. Although their study was not entirely reflective of typical clinical practice, their findings do help corroborate our findings. In our experience, pCLE allows for multiple “optical biopsies” but does not generally permit in vivo histologic analysis of the entire BE segment. Also the distal cap improves probe stabilization, but image optimization is not always feasible. These technical limitations may explain the lower sensitivity of the technology. In the one patient with HGD on random biopsy and EMR, we speculate that the pCLE probe did not contact the dysplastic area. Canto et al. demonstrated improved sensitivity of CLE compared to random biopsy using an endoscope-based version of the technology (eCLE, Pentax Medical, Montvale, NJ, USA) that included a wider surface area and provided a more stable image. Unfortunately, this technology is no longer commercially available [11].

For LGD, there was poor inter-observer agreement between pathologists as well as between physicians performing pCLE interpretation. These difficulties in making a conclusive histologic diagnosis of LGD are well documented in the literature, even among gastrointestinal pathologists with a special interest in BE [22]. In our study, the blinded pathologist used standardized criteria that optimized sensitivity, which may account for the high prevalence of LGD (29%) on blinded pathology review [19]. In routine clinical practice, pathologists may be more conservative in their assessment knowing that a diagnosis of LGD could trigger discussion with regard to multiple repeat endoscopies and ablation. Indeed, the prevalence of LGD as assessed by the unblinded pathologists was only 8%. The physician performing blinded pCLE interpretation did not know the clinical history and did not view the endoscopy images, factors that could bias real-time pCLE interpretation. For instance, an endoscopist may lean toward diagnosing dysplasia in patients with a visible nodule, ultra-long segment BE, or prior histologic diagnosis of LGD. These factors may in part account for the inter-observer variability between real-time and blinded pCLE assessments. Despite all of these limitations, both real-time and blinded pCLE demonstrated high specificity (>85%) for the diagnosis of LGD. Additionally, real time pCLE identified LGD in 75% of patients in whom both blinded and unblinded pathologists agreed on the diagnosis. Given the known limitations of histology, patients referred to tertiary referral centers for ablation of LGD frequently undergo a repeat diagnostic endoscopy with biopsies for expert pathologist review. Confirming a diagnosis of LGD with real-time pCLE at the time of repeat endoscopy could increase confidence in the diagnosis, and permit ablation during the same session.

Strengths of the study include its prospective design with a priori sample size calculations, real-time and blinded expert review of pCLE sequences, and interpretation of pathology specimens by unblinded pathologists as well as a blinded expert pathologist. Unlike previously published studies, our goal was to examine the use of pCLE in routine clinical practice (i.e., comparative effectiveness). To our knowledge, this is the first study that attempts to differentiate pCLE findings of LGD from HGD. A limitation of our study is that investigators were not required to strictly adhere to Miami criteria, because the aim was to assess accuracy in routine clinical practice. We did not use validated criteria to distinguish HGD from LGD during real-time pCLE interpretation, and there was significant disagreement between blinded and unblinded pathologists. Although inclusion of subjects was limited to a single tertiary Veterans Affairs (VA) medical center, the demographics of these patients closely resemble those of BE patients in the community setting. We did not collect information to calculate “per optical biopsy” accuracy because our aim was to assess “per patient accuracy” as suggested by the ASGE PIVI. Only three patients in our study had HGD or cancer, which has implications for estimating predictive values of pCLE. However, these findings were well within our sample size estimates, and highlight the cost-effectiveness barriers that any imaging technology faces when used for routine BE surveillance.

In summary, our study demonstrates a high specificity for dysplasia and cancer using pCLE. The relatively low sensitivity and lack of incremental benefit over HD-WLE and NBI may limit its utility in routine surveillance of BE. The technology may have a more limited role for real-time confirmation of LGD, but further study is needed to validate pCLE for this specific indication.

Competing interests
None

References

[1] Shaheen NJ, Falk NJ, Iyer PG et al. ACG Clinical Guideline: Diagnosis and Management of Barrett’s Esophagus. Am J Gastroenterol 2016; 111: 30–50
[2] Falk GW, Rice TW, Goldblum JR et al. Jumbo biopsy forceps protocol still misses unsuspected cancer in Barrett’s esophagus with high-grade dysplasia. Gastrointest Endosc 1999; 49: 170–176 PMID 9925694
[3] Kariv R, Plesec TP, Goldblum JR et al. The Seattle protocol does not more reliably predict the detection of cancer at the time of esophagectomy than a less intensive surveillance protocol. Clin Gastroenterol Hepatol 2009; 7: 653–658 quiz 606
[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: 897–909 e4; quiz 983 e1, 983 e3
[5] Yousef F, Cartwell C, Cantwell MM et al. The incidence of esophageal cancer and high-grade dysplasia in Barrett’s esophagus: a systematic review and meta-analysis. Am J Epidemiol 2008; 168: 237–249
Leggett CL, Gorospe EC. Application of confocal laser endomicroscopy in the diagnosis and management of Barrett’s esophagus. Ann Gastroenterol 2014; 27: 193 – 199

ASGE Technology Committee. Confocal laser endomicroscopy. Gastrointest Endosc 2014; 80: 928 – 938

ASGS review of Confocal Laser Endomicroscopy. 2016: Available from: https://theasgs.org/position-statements/position-statement-on-confocal-laser-endomicroscopy/

Sharma P, Brill J, Canto M et al. White Paper AGA: Advanced Imaging in Barrett’s Esophagus. Clin Gastroenterol Hepatol 2015; 13: 2209 – 2218

Bajbouj M, Vieth M, Rosch T et al. Probe-based confocal laser endomicroscopy compared with standard four-quadrant biopsy for evaluation of neoplasia in Barrett’s esophagus. Endoscopy 2010; 42: 435 – 440

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; 79: 211 – 221

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; 74: 465 – 472

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

Sharma P, Savides TJ, Canto MI et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on Imaging in Barrett’s Esophagus. Gastrointest Endosc 2012; 76: 252 – 254

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 – 166

Wallace M, Lauwers CY, Chen Y et al. Miami classification for probe-based confocal laser endomicroscopy. Endoscopy 2011; 43: 882 – 891

Spechler SJ, Sharma P, Souza RF et al. American Gastroenterological Association technical review on the management of Barrett’s esophagus. Gastroenterology 2011; 140: e18 – 52; quiz e13

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; 106: 1961 – 1969

Odze RD. Diagnosis and grading of dysplasia in Barrett’s oesophagus. J Clin Pathol 2006; 59: 1029 – 1038

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159 – 174

Abrams JA, Kapel RC, Lindberg JM et al. Adherence to biopsy guidelines for Barrett’s esophagus surveillance in the community setting in the United States. Clin Gastroenterol Hepatol 2009; 7: 736 – 742; quiz 710

Wani S, Rubenstein JH, Vieth M et al. Diagnosis and Management of Low-Grade Dysplasia in Barrett’s Esophagus: Expert Review From the Clinical Practice Updates Committee of the American Gastroenterological Association. Gastroenterology 2016; 151: 822 – 835