



Diagnostic performance of volumetric laser endomicroscopy for Barrett's esophagus dysplasia amongst gastroenterology trainees

Muhammad Aziz¹, Chandra S. Dasari², Tarun Rai², Benjamin Alsop^{2,3}, Neil Gupta⁴, Prashanth Vennalaganti², Viveksandeep Thoguluva Chandrasekar^{2,3}, Kelsey Able^{2,3}, Kevin Kennedy², Michael B. Wallace⁵, Kenneth K. Wang⁶, Herbert C. Wolfsen⁵, Prateek Sharma^{2,3}, Cadman L. Leggett⁶

¹Department of Internal Medicine, University of Toledo Medical Center, Toledo, OH, USA; ²Gastroenterology and Hepatology, Veterans Affairs Medical Center, Kansas City, MO, USA; ³Gastroenterology and Hepatology, University of Kansas Medical Center, Kansas City, KS, USA; ⁴Gastroenterology, Loyola University, Chicago, IL, USA; ⁵Gastroenterology and Hepatology, Mayo Clinic, Jacksonville, FL, USA; ⁶Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA

Contributions: (I) Conception and design: P Sharma, CL Leggett; (II) Administrative support: P Sharma; (III) Provision of study materials or patients: MB Wallace, KK Wang, HC Wolfsen, P Sharma, CL Leggett; (IV) Collection and assembly of data: B Alsop, K Able, T Rai, M Aziz; (V) Data analysis and interpretation: K Kennedy, M Aziz; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Muhammad Aziz, MD. Department of Internal Medicine, University of Toledo Medical Center, 3000 Arlington Avenue, Toledo, OH 43614, USA. Email: marajani@hotmail.com.

Background: Volumetric laser endomicroscopy (VLE) is an advanced imaging modality used in Barrett's esophagus (BE) surveillance. VLE image interpretation is challenged by subtle grayscale image variation across a large amount of data. Training in VLE interpretation is not standardized. This study aims to determine if VLE training can be incorporated into a gastroenterology (GI) fellowship curriculum with the use of a self-directed module.

Methods: A standardized, self-directed training module (30 min) was created explaining the background and established VLE criteria for the diagnosis of BE dysplasia. A VLE image dataset was generated from a multicenter VLE database of targeted biopsies. GI trainees were asked to grade each image for the presence or absence of the following criteria (I) increased surface optical frequency domain imaging (OFDI) signal intensity and (II) atypical glands and provide a final diagnosis (dysplastic *vs.* non-dysplastic). Diagnostic performance was calculated and results compared to VLE expert interpretation using histology as the gold-standard.

Results: The dataset included 50 VLE images (10 high-grade dysplasia, 40 non-dysplastic BE). VLE images were reviewed in a randomized and blinded fashion by 5 GI trainees with no prior VLE experience and 5 experienced VLE users. Sensitivity, specificity and accuracy of GI trainees was 83.3% (95% CI: 71.5–91.7%), 59.0% (95% CI: 51.6–66.0%), and 64.8% (95% CI: 58.5–70.7%) compared to 80.0% (95% CI: 67.7–89.2%), 79.5% (95% CI: 73.0–85.0%), and 79.6% (95% CI: 74.1–84.4%) for VLE experts respectively. The difference in specificity and accuracy between the two groups were statistically significant with $P < 0.001$.

Conclusions: A brief training session on VLE is inadequate to reach competency in interpretation of VLE by GI trainees. Additional experience is required to accurately interpret VLE images.

Keywords: Volumetric laser endomicroscopy (VLE); Barrett's esophagus (BE); high grade dysplasia; gastroenterology trainees (GI trainees); experts

Received: 31 October 2019; Accepted: 10 February 2020; Published: 25 January 2022.

doi: 10.21037/tgh.2020.02.15

View this article at: <http://dx.doi.org/10.21037/tgh.2020.02.15>

Introduction

Barrett's esophagus (BE) is the precursor lesion to esophageal adenocarcinoma with a proportional annual risk of progression to the highest grade of dysplasia found on surveillance endoscopy (1). Endoscopic surveillance consists of targeted biopsies (or resection) to visible lesions followed by four quadrant random biopsies every 1–2 cm along the length of BE segment (2). This practice is time consuming and prone to sampling error. Volumetric laser endomicroscopy (VLE) is an advanced imaging modality that is capable of wide-field cross-sectional imaging at microscopic resolution that may help enhance detection of dysplasia by overcoming limitations associated with surveillance endoscopy (3). VLE features associated with BE dysplasia include the presence of atypical epithelial glands and increased optical frequency domain imaging (OFDI) signal surface intensity (4–6). These features can be difficult to recognize given the need for interpretation of subtle grayscale image variation across large scale data (1,200 frames). Experienced users are capable of using VLE to distinguish between gastroesophageal tissue types including BE with and without dysplasia (6). Trindade *et al.* studied the learning curve associated with VLE interpretation amongst novice users and showed that 71% of users achieve competency after review of 96 images (7). A recent study by Alshelleh *et al.* demonstrated that patients who underwent VLE (with or without laser marking) had a higher dysplasia and neoplasia detection yield compared to traditional Seattle biopsy protocol (8). Smith *et al.* confirmed these results in a large patient population from multiple centers and validated the importance of VLE in assisting the endoscopist in tissue acquisition and treatment (9).

The aim of our study was to determine if VLE interpretation can be incorporated into a gastroenterology (GI) fellowship teaching curriculum through a standardized training module. We compare the diagnostic performance of novice users (GI fellows) following VLE training to experts in VLE interpretation using a standardized VLE image dataset.

Methods

VLE dataset

This study uses VLE images from a prospectively collected VLE database obtained at three different centers (Mayo Clinic, Rochester, MN; Mayo Clinic, Jacksonville, FL; VA Medical Center, Kansas City, MO) using the *NVision* VLE

Imaging System (NinePoint Medical Inc., Bedford, MA, USA) as part of an IRB approved study at each site. VLE was performed in patients with an established diagnosis of BE who provided written consent for participation. Patients with established contraindications to surveillance endoscopy and/or VLE imaging were excluded from the study. These images were utilized in a previous study (10).

Endoscopic surveillance

Patients enrolled in the study underwent high-definition white-light endoscopy with narrow-band imaging followed by VLE. This study used a prior version of VLE without laser marking capability. Consequently, targeted biopsies from VLE areas were obtained by triangulating the quadrant orientation and distance from incisors. Random biopsies were then obtained following a standard surveillance protocol. Histopathology was reviewed by pathologists with experience in BE interpretation (G.T., S.M., and R.C.) with highest grade of dysplasia reported per target as non-dysplastic BE (NDBE), low-grade dysplasia (LGD), high-grade dysplasia (HGD) and intramucosal adenocarcinoma. Patients with a targeted biopsy site diagnosis of NDBE or HGD were selected for the purpose of this study. LGD targets were excluded from our dataset given that VLE criteria for LGD have not been established (*Figure 1*).

VLE image selection

A high-quality full VLE scan was defined as a 6 cm scan that showed proximal stomach, gastroesophageal junction and distal esophagus with appropriate image balloon centering. Targeted biopsy sites were located on VLE scans using longitudinal distance (cm from the incisors) and angular position (o'clock position). A single VLE image was selected corresponding to each targeted biopsy site.

VLE training module

A standardized, self-directed training module was developed that explained the background of VLE and demonstrated VLE criteria with detailed explanations and example images. The training module consisted of a 30-minute video followed by a detailed discussion of the VLE features including OFDI signal intensity and atypical glandular structures. A total of 30 images with diagnosis and supporting features labelled with arrows were included

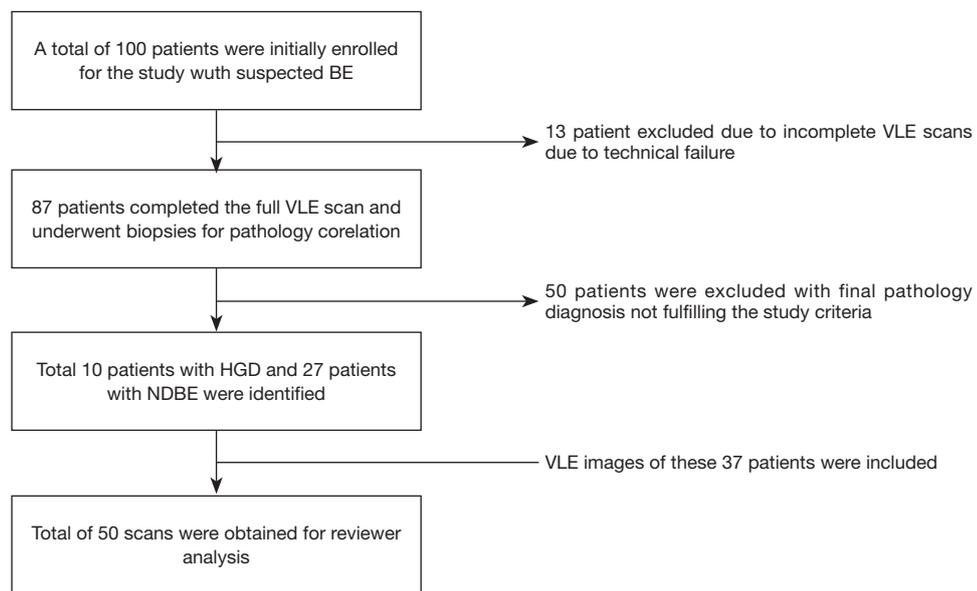


Figure 1 Selection of VLE images for reviewer analysis. BE, Barrett's esophagus; VLE, volumetric laser endomicroscopy; HGD, high-grade dysplasia; NDBE, non-dysplastic Barrett's esophagus.

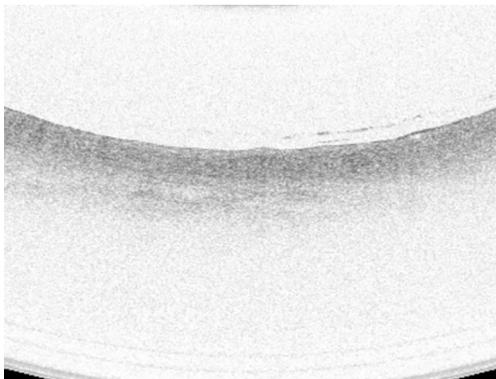


Figure 2 VLE image of NDBE demonstrating deeper OFDI signal. Also notice lack of submucosal structure. VLE, volumetric laser endomicroscopy; NDBE, non-dysplastic Barrett's esophagus; OFDI, optical frequency domain imaging.

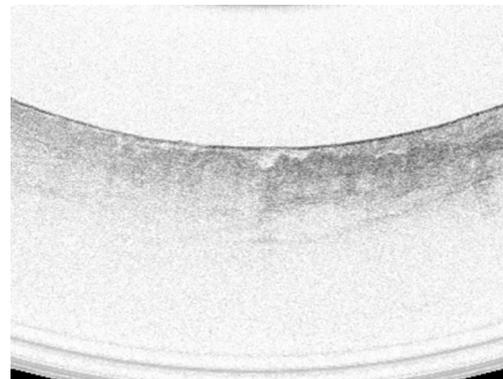


Figure 3 VLE image of HGD demonstrating more superficial OFDI signals, distorted architecture, and heterogeneous epithelium. VLE, volumetric laser endomicroscopy; HGD, high-grade dysplasia; OFDI, optical frequency domain imaging.

as part of this video (*Figures 2,3*). These images were later excluded for the main study. Each reviewer (fellow and expert) went through the tutorial video at their convenience and were given the opportunity to ask clarifying questions prior to rating the VLE dataset.

VLE assessment

A dataset of 50 high-quality VLE images was obtained from

37 patients including 10 targets of HGD and 27 targets of NDBE. The images were correlated to histopathology based on the biopsies obtained after the VLE procedures were performed. The dataset was randomized and study participants (fellows) were blinded to histology before review. All GI trainees had no previous experience of using VLE but were familiar with the use of NBI (since each site had Olympus endoscopes). The study was offered as voluntary participation and those fellows who were willing,

participated in the assessment. A panel of blinded VLE experts (Prateek Sharma, Cadman L. Leggett, Michael B. Wallace, Guillermo J. Tearney, and Herbert C. Wolfsen) was selected to serve as a comparison group. Each VLE expert had experience of reading at least 100 scans prior to this study. Reviewers evaluated the images on high-definition computer monitors and recorded the answers in a standardized table. Reviewers were asked to grade each image for: (I) presence or absence of increase OFDI surface signal intensity and (II) presence or absence of atypical glands. Reviewers were also asked to provide a diagnosis of dysplastic vs. non-dysplastic BE for each target.

Data analysis

The response from each reviewer was extracted into Microsoft Excel 2013. Statistical Analysis was performed using SPSS v23 (IBM, Armonk, NY). Sensitivity, specificity and accuracy were calculated for GI trainee and VLE expert

groups. Interobserver agreement was calculated using kappa statistics. The kappa values were graded based on the Landis and Koch scale (kappa values 0.01–0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; >0.80, nearly perfect agreement). Since there were no previous studies evaluating VLE criteria, formal sample size estimation was not performed.

Results

Patient demographics obtained from previous study are briefly summarized in *Table 1*. A total of 50 scans (10 HGD and 27 NDBE) were identified from a total of 37 patients and included in the study. The maximal extend of the BE segment was 2.7 cm (*Table 1*).

Diagnostic performance of GI trainees

A total of 5 GI trainees and 5 VLE experts participated in the rating of the VLE images. The results of the study are summarized in *Table 2*. The sensitivity, specificity and accuracy were 83.3% (95% CI: 71.5–91.7%), 59.0% (95% CI: 51.6–66.0%) and 64.8% (95% CI: 58.5–70.7%) respectively for GI trainees. For experts the sensitivity, specificity and accuracy were 80.0% (95% CI: 67.7–89.2%), 79.5% (95% CI: 73.0–85.0%) and 79.6% (95% CI: 74.1–84.4%) respectively. The difference in specificity and accuracy between the two groups were statistically significant ($P < 0.001$). The inter-observer agreement was 0.30 (95% CI: 0.21–0.40), 0.51 (95% CI: 0.40–0.63), and 0.40 (95% CI: 0.32–0.48) for GI trainees, experts and combined respectively. The difference between the inter-observer agreement between the experts and trainees was statistically significant ($P = 0.006$).

Results on the basis of individual feature are presented in *Tables 3–5*. The individual feature performed better in the experts group compared to trainees.

Table 1 Baseline demographic and EGD findings on patients undergoing the study

Characteristic	Mean (%)	Range	95% CI
Age (years) (N=100)	66.4	42, 92	64.4, 68.3
Gender (N=100)			
Male	90 (90.0)		
Female	10 (10.0)		
Race (N=100)			
Black or African American	1 (1.0)		
Caucasian	98 (98.0)		
Unknown or not reported	1 (1.0)		
Prague classification (N=100)			
C (cm)	1.4	0, 14	0.8, 2.0
M (cm)	2.7	0, 14	2.1, 3.4

EGD, esophagogastroduodenoscopy.

Table 2 Performance characteristics of trainees and experts

Characteristics	Trainees	Experts	Overall	P value
Accuracy (%)	64.8 (58.5, 70.7)	79.6 (74.1, 84.4)	72.2 (68.1, 76.1)	<0.001
Specificity (%)	59.0 (51.6, 66.0)	79.5 (73.0, 85.0)	69.2 (64.3, 73.8)	<0.001
Sensitivity (%)	83.3 (71.5, 91.7)	80.0 (67.7, 89.2)	81.7 (73.6, 86.1)	0.34
Kappa	0.30 (0.21, 0.40)	0.51 (0.40, 0.63)	0.40 (0.32, 0.48)	0.006

Table 3 Performance of individual criterion among fellows

Criterion	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	Kappa
OFDI signal (darker in superficial)	78.3 (65.8, 87.9)	52.1 (44.8, 59.4)	58.4 (52.0, 64.6)	34.0 (29.7, 38.7)	88.4 (82.2, 92.6)	0.21 (0.12, 0.31)
Architecture of acini (irregular)	68.3 (55.0, 79.7)	78.4 (71.9, 84.1)	76.0 (70.2, 81.2)	50 (42.0, 58.0)	88.7 (84.3, 91.9)	0.42 (0.29, 0.54)

OFDI, optical frequency domain imaging; PPV, positive predictive value; NPV, negative predictive value.

Table 4 Performance of individual criterion among experts

Criterion	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	Kappa
OFDI signal (darker in superficial)	73.3 (60.3, 83.9)	74.7 (67.9, 80.8)	74.4 (68.5, 79.7)	47.8 (40.7, 55.0)	89.9 (85.3, 93.2)	0.41 (0.29, 0.52)
Architecture of acini (irregular)	76.7 (64.0, 86.6)	84.7 (78.8, 89.5)	82.8 (77.5, 87.3)	61.3 (52.5, 69.5)	92 (87.9, 94.8)	0.57 (0.45, 0.68)

OFDI, optical frequency domain imaging; PPV, positive predictive value; NPV, negative predictive value.

Table 5 Performance of individual criterion among experts and fellows

Criterion	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	Kappa
OFDI signal (darker in superficial)	75.8 (67.2, 83.2)	63.4 (58.4, 68.3)	66.4 (62.1, 70.5)	39.6 (35.7, 43.6)	89.3 (85.7, 92.0)	0.30 (0.22, 0.37)
Architecture of acini (irregular)	72.5 (63.6, 80.3)	81.6 (77.3, 85.4)	79.4 (75.6, 82.9)	55.4 (49.5, 61.2)	90.4 (87.5, 92.7)	0.49 (0.41, 0.57)

OFDI, optical frequency domain imaging; PPV, positive predictive value; NPV, negative predictive value.

Discussion

VLE is a clinically available advanced imaging modality used in BE surveillance to enhance detection of dysplasia (3,11). VLE interpretation can be challenging due to subtle variation in grayscale images across a large amount of data. Training in VLE is not standardized and this study aimed to identify whether a training module can be incorporated into a GI fellowship curriculum.

Established VLE features of BE dysplasia include increase surface OFDI signal intensity and atypical epithelial glands (4-6). GI trainees were asked to identify these VLE features in an image dataset following a self-directed training module. A previous study on the learning curve of VLE showed a favorable learning curve for novice users interpreting VLE images. Approximately, 50% of the users achieved competency at 65 images and 71% achieved interpretation competency upon review of 96 images (7). Our training module contained 30 images that were reviewed prior to undergoing assessment. The duration of

the training was selected arbitrarily for this study. However, there is good evidence from previous studies on advanced imaging (such as NBI) that a 30-minute training period is adequate (12). We demonstrate that GI trainees who reviewed this brief video-format training module were capable of achieving 64.8% accuracy in the diagnosis of dysplasia compared to a 79.6% accuracy for experienced VLE users. This suggests that a brief self-directed training module is insufficient for novice VLE users to reach competency. However, experienced VLE users also demonstrate a suboptimal performance highlighting that VLE image interpretation is challenging. Moreover, all reviewers were blinded to any information related to VLE images which may have caused them to pick dysplasia due to overwhelming number of non-dysplastic images. The study by Trindade *et al.* is more practical and carries more potential to be employed in a teaching program for novice users to achieve competency (7).

Several study limitations may explain these findings. Previous literature on VLE reported several criteria

such as Evans, Leggett, and Amsterdam to differentiate between dysplastic and non-dysplastic BE. Our study is limited as we did not use any of the complete criteria rather components of different criteria. We also did not use 'effacement' which may have contributed to overall suboptimal outcomes. The VLE dataset were reviewed as limited sections and comprised of single-frame images. VLE images were reviewed using a high-definition computer monitor as compared to the designated VLE console which may have had an impact of the diagnostic performance of reviewers. Reviewers underwent self-directed training which can result in different levels of understanding and dissimilar competency in evaluating the criteria on VLE images. Not all VLE experts were actively reading VLE images at the time of study. Including active users with experience as experts would have certainly widened the gap in the results. Lastly, biopsies were taken after the VLE examination and there is always a chance that the images do not directly correlate with the biopsies that were obtained. VLE images were obtained using the older generation NinePoint model which did not have the laser marking capability. However, the longitudinal distance and angular position of the biopsy targets were used to establish imaging-to-histology correlation.

VLE image enhancement with intelligent real-time image segmentation (IRIS) is a promising tool to help in the interpretation of VLE. IRIS is an artificial intelligence-based system used to identify and display VLE features of dysplasia using a color graded scale. This software assists the VLE users to interpret VLE images and identify a region of interest for targeted biopsies (13). The use of IRIS coupled with long term training for novice users may lead to better assessment and interpretation of VLE images with higher confidence. Further, the use of laser marking with new generation VLE can also increase the dysplasia and neoplasia detection rates.

In conclusion, the difference in performance between experts and trainees suggests that a brief training session on VLE is inadequate. Although GI trainees were able to interpret VLE images after a self-directed teaching session, perhaps both the duration and type of teaching are inadequate to them to reach levels of accuracy compared to VLE experts (7). Future studies involving computer-assisted detection should explore approaches for improving education for novel imaging techniques.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://tgh.amegroups.com/article/view/10.21037/tgh.2020.02.15/coif>). Dr. MBW reports grants from Olympus Inc, grants from Boston Scientific, grants from NinePoint Medical, outside the submitted work. Dr. KKW reports research support from NinePoint Medical, CSA Medical, Fujinon, CDx Diagnostics, outside the submitted work. Dr. PS reports grants from US Endoscopy, grants from Medtronic, grants from Fujifilm, grants from Ironwood, grants from Cosmo pharmaceuticals, grants from Erbe, personal fees from Boston Scientific, personal fees from Olympus Inc, outside the submitted work. Dr. CLL reports non-financial support from NinePoint Medical, during the conduct of the study. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The current study utilized de-identified data and images from a previously conducted multi-center study (Wolfsen *et al.*) that had ethics approval at each site and hence another ethics/IRB approval was not obtained. VLE was performed in patients with an established diagnosis of BE who provided written consent for participation.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Bhat S, Coleman HG, Yousef F, et al. Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study. *J Natl Cancer Inst* 2011;103:1049-57.
2. 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.
3. Aziz M, Fatima R. Future of diagnosing neoplasia in Barrett's esophagus: volumetric laser endomicroscopy. *Clin J Gastroenterol* 2018;11:179-83.
 4. Evans JA, Ponerros JM, Bouma BE, et al. Optical coherence tomography to identify intramucosal carcinoma and high-grade dysplasia in Barrett's esophagus. *Clin Gastroenterol Hepatol* 2006;4:38-43.
 5. Swager AF, Tearney GJ, Leggett CL, et al. Identification of volumetric laser endomicroscopy features predictive for early neoplasia in Barrett's esophagus using high-quality histological correlation. *Gastrointest Endosc* 2017;85:918-926.e7.
 6. Leggett CL, Gorospe EC, Chan DK, et al. Comparative diagnostic performance of volumetric laser endomicroscopy and confocal laser endomicroscopy in the detection of dysplasia associated with Barrett's esophagus. *Gastrointest Endosc* 2016;83:880-888.e2.
 7. Trindade AJ, Inamdar S, Smith MS, et al. Learning curve and competence for volumetric laser endomicroscopy in Barrett's esophagus using cumulative sum analysis. *Endoscopy* 2018;50:471-8.
 8. Alshelleh M, Inamdar S, McKinley M, et al. Incremental yield of dysplasia detection in Barrett's esophagus using volumetric laser endomicroscopy with and without laser marking compared with a standardized random biopsy protocol. *Gastrointest Endosc* 2018;88:35-42.
 9. Smith MS, Cash B, Konda V, et al. Volumetric laser endomicroscopy and its application to Barrett's esophagus: results from a 1,000 patient registry. *Dis Esophagus* 2019;32:doz029.
 10. Wolfsen HC, Sharma P, Wallace MB, et al. Safety and feasibility of volumetric laser endomicroscopy in patients with Barrett's esophagus (with videos). *Gastrointest Endosc* 2015;82:631-40.
 11. Trindade AJ, Smith MS, Pleskow DK. The new kid on the block for advanced imaging in Barrett's esophagus: a review of volumetric laser endomicroscopy. *Therap Adv Gastroenterol* 2016;9:408-16.
 12. Raghavendra M, Hewett DG, Rex DK. Differentiating adenomas from hyperplastic colorectal polyps: narrow-band imaging can be learned in 20 minutes. *Gastrointest Endosc* 2010;72:572-6.
 13. Volumetric Laser Endomicroscopy With Intelligent Real-time Image Segmentation (IRIS) - Full Text View - ClinicalTrials.gov. *Clinicaltrials.gov*. Available online: <https://clinicaltrials.gov/ct2/show/NCT03814824>. Published 2019. Accessed March 8, 2019.

doi: 10.21037/tgh.2020.02.15

Cite this article as: Aziz M, Dasari CS, Rai T, Alsop B, Gupta N, Vennalaganti P, Chandrasekar VT, Able K, Kennedy K, Wallace MB, Wang KK, Wolfsen HC, Sharma P, Leggett CL. Diagnostic performance of volumetric laser endomicroscopy for Barrett's esophagus dysplasia amongst gastroenterology trainees. *Transl Gastroenterol Hepatol* 2022;7:3.