

# Expert-led didactic versus self-directed audiovisual training of confocal laser endomicroscopy in evaluation of mucosal barrier defects



## Authors

Roy Huynh<sup>1,2</sup>, Matthew Ip<sup>1,2</sup>, Jeff Chang<sup>1,2,3</sup>, Craig Haifer<sup>3</sup>, Rupert W. Leong<sup>1,2,3,4</sup>

## Institutions

- 1 Faculty of Medicine, UNSW Australia, Sydney, Australia
- 2 Gastroenterology and Liver Services, Bankstown-Lidcombe Hospital, Sydney, Australia
- 3 Gastroenterology and Liver Services, Concord Repatriation General Hospital, Sydney, Australia
- 4 Australian School of Advanced Medicine, Macquarie University, Sydney, Australia

submitted 28.1.2017

accepted after revision 14.6.2017

## Bibliography

DOI <https://doi.org/10.1055/s-0043-114664> |

Endoscopy International Open 2018; 06: E115–E122

© Georg Thieme Verlag KG Stuttgart · New York

ISSN 2364-3722

## Corresponding author

Professor Rupert Leong, Gastroenterology and Liver Services, Bankstown-Lidcombe and Concord Hospital, c/o Concord Hospital Level 1 West, Hospital Rd, Concord NSW 2137, Sydney, Australia  
 Fax: +61 2 97227752  
[rupertleong@outlook.com](mailto:rupertleong@outlook.com)

## ABSTRACT

**Background and study aims** Confocal laser endomicroscopy (CLE) allows mucosal barrier defects along the intestinal epithelium to be visualized *in vivo* during endoscopy.

Training in CLE interpretation can be achieved didactically or through self-directed learning. This study aimed to compare the effectiveness of expert-led didactic with self-directed audiovisual teaching for training inexperienced analysts on how to recognize mucosal barrier defects on endoscope-based CLE (eCLE).

**Materials and methods** This randomized controlled study involved trainee analysts who were taught how to recognize mucosal barrier defects on eCLE either didactically or through an audiovisual clip. After being trained, they evaluated 6 sets of 30 images. Image evaluation required the trainees to determine whether specific features of barrier dysfunction were present or not. Trainees in the didactic group engaged in peer discussion and received feedback after each set while this did not happen in the self-directed group. Accuracy, sensitivity, and specificity of both groups were compared.

**Results** Trainees in the didactic group achieved a higher overall accuracy (87.5% vs 85.0%,  $P=0.002$ ) and sensitivity (84.5% vs 80.4%,  $P=0.002$ ) compared to trainees in the self-directed group. Interobserver agreement was higher in the didactic group ( $k=0.686$ , 95% CI 0.680–0.691,  $P<0.001$ ) than in the self-directed group ( $k=0.566$ , 95% CI 0.559–0.573,  $P<0.001$ ). Confidence (OR 6.48, 95% CI 5.35–7.84,  $P<0.001$ ) and good image quality (OR 2.58, 95% CI 2.17–2.82,  $P<0.001$ ) were positive predictors of accuracy.

**Conclusion** Expert-led didactic training is more effective than self-directed audiovisual training for teaching inexperienced analysts how to recognize mucosal barrier defects on eCLE.

## Introduction

Confocal laser endomicroscopy (CLE) allows real-time magnification of the intestinal epithelium for the *in vivo* detection of mucosal barrier defects [1–3]. The learning curve for CLE has previously been demonstrated; inexperienced analysts, trained with a short didactic session, can reach a level of accuracy equivalent to that seen in experts after exposure to only 90 images [4]. However, optimal training methods remain unclear.

In order for CLE to be widely utilized, evidence-based programs must be developed to allow standardized training for clinicians and researchers [5, 6].

Current training in CLE is typically done didactically or through online resources [1]. Expert-led didactic teaching is capable of producing rapid learning curves for the interpretation of many gastrointestinal diseases on CLE [7–11]. However, it is limited by accessibility and cost. Not all endoscopy centers will have access to an expert CLE analyst and attending training

seminars can be expensive. To address this problem, a collection of online resources comprising of audiovisual clips, image atlases, and smartphone applications are available for trainees who are interested in acquiring CLE interpretation skills [1]. Trainee analysts who use these resources have flexibility over content, learning pace, and order of learning [12]. Internet distribution of educational content has made it possible to acquire CLE interpretation skills in the absence of an expert CLE analyst. Whether these resources can be utilized to the same effect as expert-led didactic teaching remains debatable. Previous studies analyzing learning curves and pedagogics of CLE have yet to compare training methods for evaluation of mucosal barrier defects on eCLE [7–16].

Didactic teaching is the most frequently employed training method in CLE studies [7–11]. Effective didactic teaching involves small class sizes to optimize interaction between trainees and encourages active participation to promote long-term retention [17]. The instructor must be capable of addressing individual learning needs and flexible with a range of different teaching styles. In contrast self-directed learning require trainees to take complete responsibility for their education through self-accessible resources [18]. Given the challenge of increasing healthcare delivery and less teaching time, medical educators are now increasingly turning over control of learning to the students [19,20]. Audiovisual clips are a popular way to deliver CLE lessons and can be paused and rewound for content control [12]. Self-directed audiovisual learning may represent a more cost-effective pedagogic approach to didacticism for training CLE analysts.

The primary aim of this study was to compare the effectiveness of expert-led didactic with self-directed audiovisual teaching for training inexperienced analysts how to recognize mucosal barrier defects on eCLE. Qualitative feedback was obtained from trainees in both groups to establish the advantages and limitations of each pedagogic approach. The secondary aim was to confirm previous findings that image quality and analyst confidence are positive predictors of accuracy [4].

## Materials and methods

### Image collection

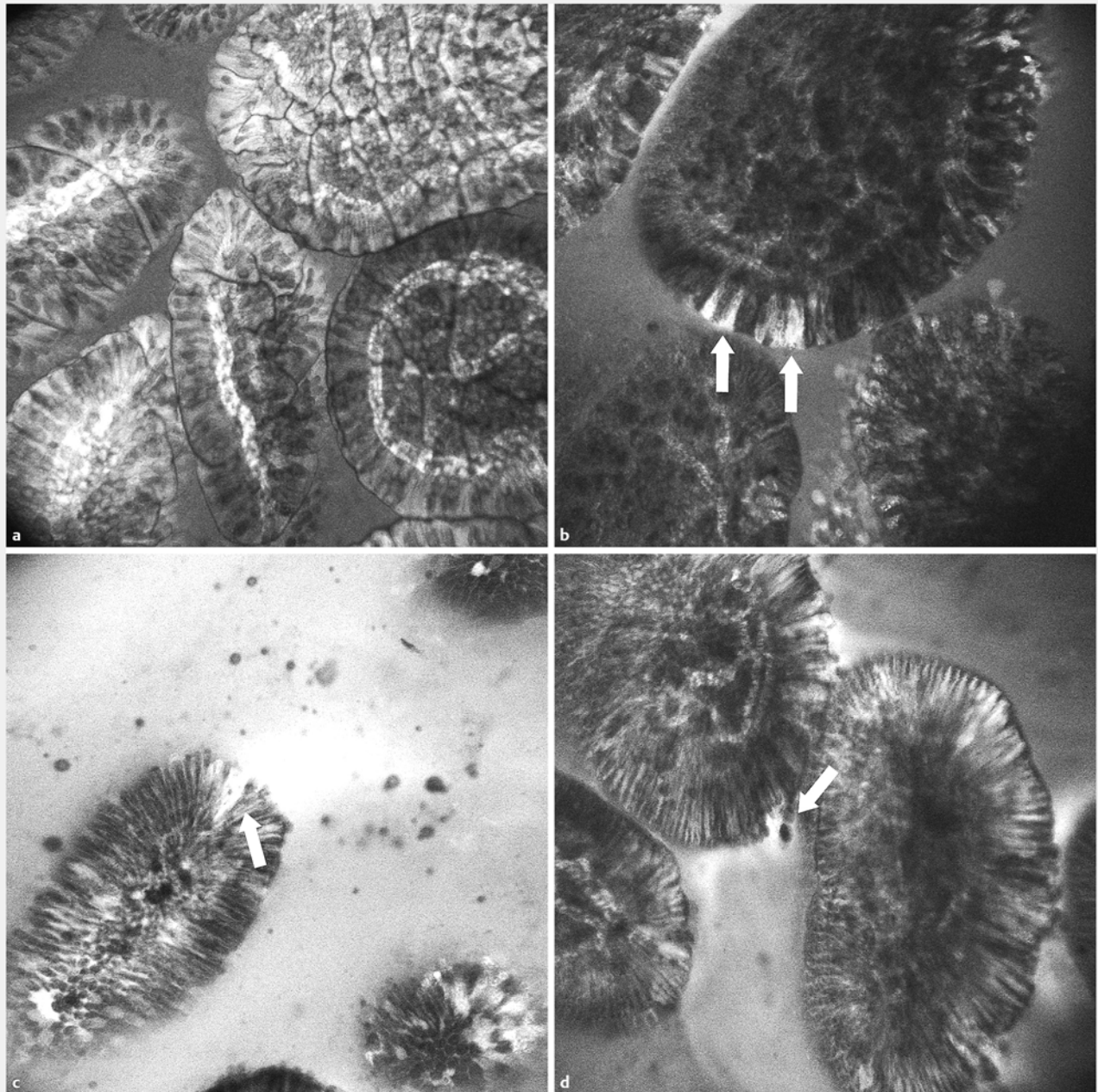
180 eCLE images were selected from our database to be included in this study. The images were taken from patients who required a colonoscopy for clinical indications. Informed consent was obtained from patients over age 18 who did not fulfill the exclusion criteria of pregnancy, advanced kidney disease, or known fluorescein allergy. The procedure was carried out at Bankstown-Lidcombe Hospital, Sydney, Australia. Patients were given standard conscious sedation with propofol and 5 mL of intravenous 10% fluorescein sodium as a contrast agent. Images were taken from 5 sites of the terminal ileum with paired biopsies. The selected images were assessed by an experienced eCLE analyst as either good or poor quality. Good-quality images were defined as those that contained at least one-3<sup>rd</sup> of the visual field free from movement artefacts. This study was approved by the Human Research Ethics Committee of Sydney South Area Health (Reference number 14/184).

### Reference standard

The appearance of each mucosal barrier defect on eCLE was initially derived from benchtop intra-vital 2-photon confocal laser microscopy of explanted murine intestines. Intravenous Hoechst 33258 was used to stain the epithelial cellular nuclei of mice before 5 µg of tumor necrosis factor-α was administered to induce mucosal cell shredding for comparison with controls that were given intraperitoneal phosphate buffer solution. Under confocal laser microscopy (Carl Zeiss 7MP; Carl Zeiss, Oberkochen, Germany and Chameleon Vision II Ti:Sapphire laser; Coherent Scientific, Hilton, South Australia, Australia), compromised epithelial features were identified using the image analysis program IMARIS [21]. Pilot studies of eCLE in the human terminal ileum demonstrated that similar features could be observed in controls and patients with inflammatory bowel disease [22,23]. After disregarding features with poor interobserver agreement, a consensus was reached among CLE experts to recognize 3 main features of barrier dysfunction: cell junction enhancement (CJE), fluorescein leak (FL), and cell dropout (CDO) [22–24]. Interobserver kappa values for CJE, FL, and CDO have been demonstrated to be high among experienced analysts (0.85, 0.92, and 0.80 respectively) [4]. In CJE, the tight junction proteins are compromised, permitting fluorescein to accumulate in the intercellular junction, resulting in highlighting of adjacent enterocytes (► Fig. 1b). The appearance of FL can be appreciated as the outflow of fluorescein into the lumen due to the complete loss of apposition between adjacent enterocytes (► Fig. 1c). Shredding of an apoptotic cell, usually into the vicinity of a fluorescein plume is observed in CDO (► Fig. 1d) [24]. These validated features have been used in previous studies to identify barrier dysfunction in many gastrointestinal diseases [22–26]. Severity of barrier dysfunction can be quantified using the Confocal Leak Score (CLS), which is expressed as a percentage of the number of images displaying FL, CJE, and CDO over the total number of images analyzed [4]. CLS assesses intestinal permeability on a continuous scale, rather than a categorical variable as achieved by the Watson grading system [23,27].

### Training and image evaluation

Trainee analysts with no prior experience with eCLE were randomized by sequential group allocations to either the didactic or self-directed group. Randomization occurred using a binary number generator after each group had formed. We did not conceal as randomization occurred immediately prior to each teaching session to remove any allocation bias. Didactic sessions were conducted by a lecturer with experience interpreting over 100 eCLE cases who therefore fit the definition of an experienced eCLE analyst [4]. The same lecturer was used for all the sessions to ensure that the teaching method was consistent. A PowerPoint presentation composed of 9 eCLE training images was used to explain examples of FL, CJE, and CDO. Trainees in the didactic group were encouraged to engage in peer discussion throughout the presentation. They also had direct interaction with the lecturer which enabled them to ask questions when needed.



► **Fig. 1** eCLE images displaying **a** normal mucosa, **b** cell junction enhancement (CJE), **c** fluorescein leak (FL), and **d** cell dropout (CDO).

In contrast, trainees in the audiovisual group were completely self-directed in their learning. They were provided with an online audiovisual clip comprising of the same PowerPoint slides that were presented to the didactic group. The slides contained the lecturer's prerecorded voice overlaid on top. Trainees in this group could adjust the viewing pace or pause and rewind the clip when needed. However, they did not have the advantage of asking the lecturer questions or engage in peer discussion.

Image evaluation took place after the training and was composed of 180 eCLE images subdivided into 6 sets of 30 images. Each set contained equal numbers of images assessing FL, CJE,

and CDO. For each image, the trainees had to determine the presence or absence of an asked feature. They also had to record whether they were confident or not about each response. After each set, the correct answers were provided to the trainees for self-reflection. Trainees in the didactic group had the opportunity to review previous images with the lecturer while that was not possible in the self-directed group. The same assessment method was repeated after 6 months to determine the intraobserver agreement in both groups. An open questionnaire was used to obtain qualitative feedback from the trainees.

► **Table 1** Comparison of overall accuracy, sensitivity, and specificity.

	Expert-led didactic	Self-directed audiovisual	P value
% Accuracy (95% CI)	87.5 (86.5–88.5)	85.0 (83.9–86.2)	0.002
% Sensitivity (95% CI)	84.5 (83.1–85.9)	80.4 (78.1–82.6)	0.002
% Specificity (95% CI)	90.1 (88.7–91.5)	89.2 (87.4–91.0)	0.438

## Sample size and statistical analysis

Statistical analysis was performed using IBM SPSS 20.0 (IBM Corp, Armonk, NY, USA). The sample size was determined using preliminary data from a pilot study of the same design on 8 trainee analysts. This identified a mean difference of 7.5% (92.5% vs 85.0%, SD 0.071) between the didactic and self-directed group respectively. A minimum of 20 trainee analysts was needed in each group according to 2-proportion unpaired *t*-test using a 2-tailed  $\alpha$  of 0.05 and a power (1- $\beta$ ) of 0.90. 2-sample *t*-test was used to compare mean accuracy, sensitivity, and specificity between the 2 groups. Multilevel logistic regression was performed using generalized estimating equations to determine the impact image quality and confidence has on accuracy. Fleiss' and Cohen's kappa were used to assess interobserver and intraobserver agreement, respectively. Level of agreement was acknowledged as no agreement ( $k=0$ ), poor agreement ( $k=0.01–0.20$ ), mild agreement ( $k=0.21–0.40$ ), fair agreement ( $k=0.41–0.60$ ), moderate agreement ( $k=0.61–0.80$ ), strong agreement ( $k=0.81–0.99$ ), and perfect agreement ( $k=1$ ). Statistical significance was defined as *P* values less than 0.05.

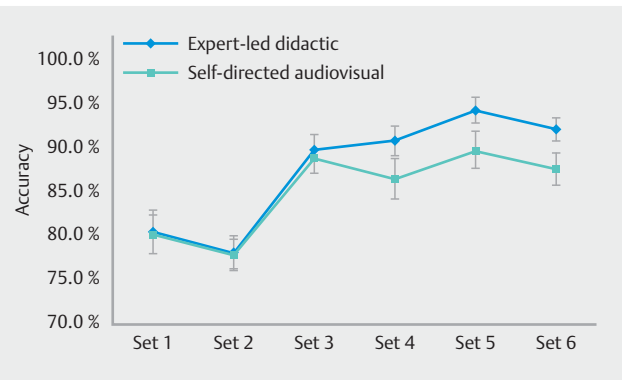
## Results

### Study participants

72 trainee analysts, all of whom were medical students participated in this study. The rationale behind choosing medical students over untrained pathologists or endoscopists was to obtain large sample size for increased statistical power. Trainee level of knowledge regarding gastrointestinal pathology has previously been demonstrated to have no impact on ability to acquire eCLE interpretation skills for mucosal barrier defects [4]. Recruitment took place at 2 metropolitan teaching hospitals in Sydney, Australia: Bankstown-Lidcombe Hospital and Concord Hospital. Randomization via sequential group allocation resulted in 36 trainees in the didactic group and also in the self-directed group. Trainees were not induced with any incentives besides the opportunity to learn a new imaging technique.

### Comparison of learning curves

Overall accuracy, sensitivity, and specificity for both groups are compared in ► **Table 1**. Trainees who were taught didactically achieved a higher overall accuracy (87.5% vs 85.0%,  $P=0.002$ ) and sensitivity (84.5% vs 80.4%,  $P=0.002$ ) compared to trainees who were self-directed. There was no difference in overall specificity between the 2 groups (90.1% vs 89.2%,  $P=0.44$ ).



► **Fig. 2** Diagnostic accuracy with 95% confidence interval: comparison of learning curves between the didactic (diamond) and self-directed groups (square).

Trainees in the didactic group displayed moderate interobserver agreement ( $k=0.686$ , 95% CI 0.680–0.691,  $P<0.001$ ), while trainees in the self-directed group exhibited fair interobserver agreement ( $k=0.566$ , 95% CI 0.559–0.573,  $P<0.001$ ).

A comparison of the learning curves between the trainees who were taught didactically and the trainees who were self-directed is shown in ► **Fig. 2**. Initially, there was no difference in accuracy between the 2 groups as seen in the 1<sup>st</sup> (80.5% vs 80.2%,  $P=0.86$ ), 2<sup>nd</sup> (78.1% vs 78.0%,  $P=0.94$ ), and 3<sup>rd</sup> image set (89.7% vs 88.5%,  $P=0.27$ ). However, trainees in the didactic group ended up achieving higher accuracy from the 4<sup>th</sup> image set (90.6% vs 86.4%,  $P=0.003$ ) onward (► **Table 2**).

Mean sensitivity achieved by each group in each image set is compared in ► **Fig. 3**, which shows similarities to the comparison of learning curves. At the outset, trainees in the didactic group were no more sensitive than the trainees in the self-directed group, as seen in the 1<sup>st</sup> (73.3% vs 74.7%,  $P=0.58$ ), 2<sup>nd</sup> (73.6% vs 69.6%,  $P=0.10$ ), and 3<sup>rd</sup> image sets (88.1% vs 85.0%,  $P=0.08$ ). But by the 4<sup>th</sup> image set, trainees in the didactic group were able to achieve a higher sensitivity than the trainees in the self-directed group (93.1% vs 82.3%,  $P<0.001$ ). This trend continued up to the 6<sup>th</sup> image set (92.8% vs 83.3%,  $P<0.001$ ) (► **Table 2**).

Although there was no difference in overall specificity between the 2 groups, trainees in the didactic group managed to achieve higher specificity in the fifth image set (98.7% vs 94.7%,  $P=0.003$ ) compared to trainees in the self-directed group. However, that difference was not sustained in the 6<sup>th</sup> image set (91.4% vs 90.3%,  $P=0.46$ ) and is therefore likely due to chance (► **Table 2**).

► **Table 2** Comparison of accuracy, sensitivity, and specificity of each set between the 2 groups

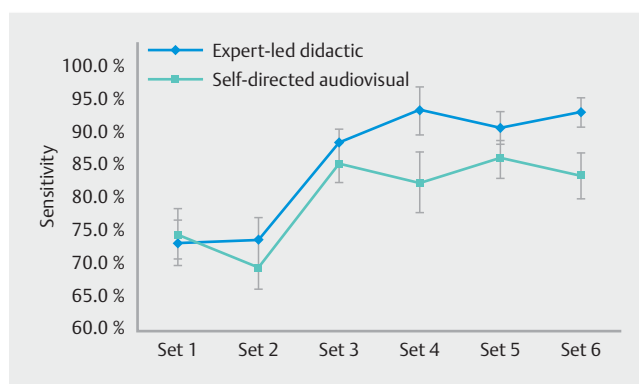
		Expert-led didactic	Self-directed audiovisual	P value
Set 1	% Accuracy (95% CI)	80.5 (78.1 – 82.8)	80.2 (78.1 – 82.3)	0.859
	% Sensitivity (95% CI)	73.3 (69.9 – 76.6)	74.7 (70.9 – 78.4)	0.576
	% Specificity (95% CI)	88.7 (85.7 – 91.7)	86.5 (83.0 – 90.1)	0.342
Set 2	% Accuracy (95% CI)	78.1 (76.1 – 80.0)	78.0 (76.3 – 79.6)	0.941
	% Sensitivity (95% CI)	73.6 (70.2 – 77.1)	69.6 (66.3 – 73.0)	0.101
	% Specificity (95% CI)	81.9 (79.1 – 84.7)	85.2 (82.4 – 88.1)	0.097
Set 3	% Accuracy (95% CI)	89.7 (88.1 – 91.3)	88.5 (87.0 – 90.0)	0.268
	% Sensitivity (95% CI)	88.1 (85.9 – 90.3)	85.0 (82.3 – 87.7)	0.076
	% Specificity (95% CI)	92.1 (90.1 – 94.1)	92.4 (90.1 – 94.6)	0.878
Set 4	% Accuracy (95% CI)	90.6 (90.0 – 92.3)	86.4 (84.1 – 88.7)	0.003
	% Sensitivity (95% CI)	93.1 (89.5 – 96.6)	82.3 (77.7 – 86.9)	<0.001
	% Specificity (95% CI)	89.6 (87.5 – 91.7)	88.0 (84.9 – 91.1)	0.381
Set 5	% Accuracy (95% CI)	94.1 (92.6 – 95.5)	89.6 (87.6 – 91.7)	0.001
	% Sensitivity (95% CI)	90.5 (88.1 – 93.0)	85.8 (82.9 – 88.7)	0.013
	% Specificity (95% CI)	98.7 (97.7 – 99.7)	94.7 (92.2 – 97.1)	0.003
Set 6	% Accuracy (95% CI)	91.9 (90.6 – 93.2)	87.5 (85.7 – 89.3)	<0.001
	% Sensitivity (95% CI)	92.8 (90.6 – 95.1)	83.3 (79.9 – 86.8)	<0.001
	% Specificity (95% CI)	91.4 (89.2 – 93.5)	90.3 (88.2 – 92.4)	0.463

## Detection of mucosal barrier defects

Overall accuracy, sensitivity, and specificity of each group in interpreting FL, CJE, and CDO is shown in ► **Table 3**. Of the 3 features, FL displayed the highest overall accuracy, sensitivity, and specificity in both groups. Trainees who were taught didactically were able to interpret FL with a higher overall accuracy (95.3% vs 92.2%,  $P<0.001$ ) and sensitivity (93.2% vs 88.9%,  $P<0.001$ ) than trainees who were self-directed. However, there was no difference in overall specificity of FL between the 2 groups (97.7% vs 95.9%,  $P=0.07$ ).

Trainees in the didactic group were also able to interpret CJE with higher overall accuracy (84.7% vs 80.0%,  $P=0.001$ ) and sensitivity (81.5% vs 74.8%,  $P=0.001$ ) compared to the trainees in the self-directed group. There was no difference in overall specificity of CJE (65.2% vs 66.9%,  $P=0.59$ ) between the 2 groups. For interpretation of CDO, neither didactic nor audiovisual training was able to produce higher overall accuracy (82.4% vs 83.0%,  $P=0.60$ ), sensitivity (75.0% vs 76.2%,  $P=0.67$ ), or specificity (85.5% vs 85.9%,  $P=0.81$ ).

Interobserver agreement for FL is better between trainees who were taught didactically ( $k=0.872$ , 95% CI 0.862–0.882,  $P<0.001$ ) than trainees who were self-directed ( $k=0.784$ , 95% CI 0.774–0.794,  $P<0.001$ ). For interpretation of CJE, moderate interobserver agreement was seen in the didactic group ( $k=0.631$ , 95% CI 0.621–0.641,  $P<0.001$ ) while only fair interobserver agreement was seen in the self-directed group ( $k=0.495$ , 95% CI 0.485–0.505,  $P<0.001$ ). There was fair interob-



► **Fig. 3** Diagnostic sensitivity with 95% confidence interval: comparison between the didactic and (diamond) self-directed groups (square).

server agreement for CDO in both the didactic ( $k=0.502$ , 95% CI 0.492–0.512,  $P<0.001$ ) and self-directed group ( $k=0.561$ , 95% CI 0.551–0.571,  $P<0.001$ ).

## Confidence and image quality

The overall accuracy, sensitivity, and specificity of confident responses in both groups are compared in ► **Table 4**. There was no significant difference in confidence levels between the trainees who were taught didactically and the trainees who were

► **Table 3** Overall accuracy, sensitivity, and specificity of each mucosal barrier defect.

	Feature	Expert-led didactic	Self-directed audiovisual	P value
% Accuracy (95% CI)	FL	95.3 (94.5–96.2)	92.2 (91.2–93.2)	<0.001
	CJE	84.7 (83.2–86.2)	80.0 (77.7–82.3)	0.001
	CDO	82.4 (80.4–84.3)	83.0 (81.4–84.6)	0.604
% Sensitivity (95% CI)	FL	93.2 (91.8–94.6)	88.9 (87.1–90.7)	<0.001
	CJE	81.5 (79.3–83.7)	74.8 (71.3–78.3)	0.001
	CDO	75.0 (71.2–78.8)	76.2 (71.8–80.7)	0.668
% Specificity (95% CI)	FL	97.7 (96.6–98.9)	95.9 (94.3–97.5)	0.069
	CJE	89.2 (87.2–91.3)	87.2 (84.2–90.3)	0.274
	CDO	85.5 (83.1–87.9)	85.9 (83.7–88.1)	0.805

► **Table 4** Comparison of overall accuracy, sensitivity, and specificity of confident responses only.

	Expert-led didactic	Self-Directed audiovisual	Pvalue
% Confident Responses (95% CI)	69.7 (65.6–73.8)	73.4 (70.2–76.6)	0.150
% Accuracy (95% CI)	94.0 (92.9–95.1)	92.6 (91.3–94.0)	0.119
% Sensitivity (95% CI)	92.3 (90.4–94.2)	89.8 (87.3–92.3)	0.104
% Specificity (95% CI)	95.4 (94.4–96.5)	94.6 (93.0–96.3)	0.403

self-directed (69.7% vs 73.4%,  $P=0.15$ ). On multilevel logistic regression, confidence was shown to be a positive predictor of accuracy (OR 6.48, 95% CI 5.35–7.84,  $P<0.001$ ). Unsurprisingly, good quality images were more likely to be interpreted correctly than normal quality images (OR 2.58, 95% CI 2.17–2.82,  $P<0.001$ ). There was moderate interobserver agreement for interpretation of good quality images in both the didactic ( $k=0.693$ , 95% CI 0.687–0.699,  $P<0.001$ ) and self-directed group ( $k=0.633$ , 95% CI 0.627–0.639,  $P<0.001$ ). However, for interpretation of normal quality images, there was only fair interobserver agreement in the didactic group ( $k=0.501$ , 95% CI 0.488–0.514,  $P<0.001$ ) and mild interobserver agreement in the self-directed group ( $k=0.375$ , 95% CI 0.362–0.388,  $P<0.001$ ).

### Intraobserver agreement

Intraobserver agreement was determined at 6 months using paired data from 4 analysts in each group. None of the analysts participated in further eCLE training or interpreted any images during the 6 months. ► **Table 5** displays the kappa score of each analyst. Trainees in the didactic group had a higher mean intraobserver agreement ( $k=0.80$ ,  $\pm 0.03$  SD) than trainees in the audiovisual group ( $k=0.68$ ,  $\pm 0.03$ ).

### Qualitative feedback

56 trainees (didactic  $n=32$ ; audiovisual  $n=24$ ) provided qualitative feedback related to the training they received. All the trainees stated that they were better at interpreting eCLE for mucosal barrier defects after participating in this study. When

► **Table 5** Intraobserver agreement in both groups.

Expert-led didactic, k	Self-directed audiovisual, k
0.85	0.70
0.78	0.71
0.79	0.63
0.78	0.67

asked what was the best feature of their training, trainees in the didactic group mentioned personalized feedback ( $n=18$ ), simple explanations ( $n=7$ ), class interaction ( $n=5$ ), and peer discussion ( $n=2$ ). In comparison, trainees in the audiovisual group stated the freedom of choosing their own learning pace ( $n=18$ ) and environment ( $n=6$ ) were the most favorable aspect of their training. Interestingly, none of trainees from the audiovisual group found the pause and rewind function useful.

All the respondents from the didactic group enjoyed asking questions and receiving personalized feedback. This allowed them to either reaffirm existing knowledge or correct any misconceived ideas they had. With regard to areas for improvement, trainees in both the didactic ( $n=11$ ) and self-directed group ( $n=8$ ) requested more eCLE images during training. A majority of respondents from the self-directed group ( $n=19$ ) found it difficult to maintain attention throughout the entire audiovisual clip due to the lack of active engagement.

## Discussion

Our results demonstrated that expert-led didactic is capable of producing higher accuracy and sensitivity than self-directed audiovisual training in the identification of mucosal barrier defects on eCLE. However, this difference is only noticeable after the 3<sup>rd</sup> image set, suggesting that the superiority of expert-led didactic over self-directed audiovisual training only becomes significant after 90 images have been interpreted. While the overall percentage difference in accuracy (2.5%) and sensitivity (4.1%) between the 2 groups was small, it was still clinically significant as it corresponded to a large number of images interpreted incorrectly. Calculation of CLS involves interpretation of several hundred eCLE images from a single patient. Hence, a small percentage difference will correspond to a large number of images. Qualitative feedback obtained at the end of the study allowed us to identify significant elements which may have affected learning in both groups.

Audiovisual clips are becoming a fast-growing tool in medical education especially in new fields such as eCLE. They are capable of reaching a large audience through online portals and allow unlimited access to educational content [28]. Information can also be updated in a timely manner to ensure that clinicians have the latest knowledge [29]. Despite these advantages, audiovisual training raises numerous constraints such as passive learning and potential boredom as this pedagogic approach limits active participation [30]. The majority of self-directed trainees in our study stated that they found it difficult to maintain focus throughout the entire audiovisual clip. Lack of interactivity is a commonly cited problem in audiovisual training as it limits the attention span of trainees, resulting in poorly comprehended concepts that can negatively impact eCLE interpretation [31,32]. In contrast, class interaction is promoted in didactic teaching to allow trainees to discuss different interpretation styles. This sharing of experiences encouraged the trainees to think critically so they can modify their technique to improve accuracy. Experimental learning over 6 sets of images allowed the trainees to eventually determine which interpretation method works best for them.

Effective didactic teaching in eCLE requires an expert analyst who is flexible with a range of different teaching styles [6,33]. Accurate assessment of each trainee's ability is important to address individual learning needs [34]. In our study, trainees in the didactic group were constantly encouraged to ask questions to allow the instructor to identify areas of learning difficulty. In comparison, trainees who were self-directed were not given the opportunity to ask questions, making it more challenging to grasp difficult concepts. They were also not provided with personalized feedback, which would have enabled them to improve on their shortcomings [35]. Trainees in the didactic group attained an overall deeper understanding of barrier dysfunction than trainees in the self-directed group as they were repetitively told to understand the physiological changes behind each mucosal barrier defect, rather than just memorize their visual appearance. This level of deeper understanding has been proven beneficial for long-term retention [35].

In the didactic group, ability to review previous images with the instructor was a significant advantage as it allowed the trainees to rapidly improve their eCLE interpretation skills, causing them to eventually outperform the self-directed group after the 3<sup>rd</sup> image set. Trainees in the self-directed group did not have this privilege but had the ability to pause and rewind audiovisual content. However, that was not found useful by any of the trainees.

Constant feedback and discussions would have caused the trainees in the didactic group to acquire similar interpretation mindsets. This may have resulted in the higher interobserver agreement seen in the didactic group. Intraobserver agreement was also higher in the didactic group compared to the self-directed group. However, because that was determined 6 months after the initial assessment, it may simply be a reflection of higher retention rates within the didactic group.

Confidence and good image quality were shown to be positive predictors of accuracy and may therefore have a role in eCLE training [4]. Confidence can be promoted in didactic teaching by offering encouraging remarks throughout the training process and motivating trainees to persist in difficult areas [17]. Interpretation of eCLE images for mucosal barrier defects should be carried out on good quality images whenever possible as they are more likely to be interpreted correctly compared to normal quality images.

Limitations need to be acknowledged. In our study, there were no more than 5 trainees in each didactic session. However, eCLE training seminars often involve the attendance of many more trainees. In such a setting, it is possible that the quality of feedback and peer discussion would be lower, which may have an impact on the overall accuracy and sensitivity seen in the didactic group. Nonetheless, we believe that didactic teaching would still be superior to self-directed learning, where there is a complete lack of interaction and engagement. Further study comparing these 2 pedagogic methods involving larger classes in the didactic group will resolve this issue. Another study limitation is that post-procedural images were used instead of live images at the time of eCLE. However, that was necessary to provide the same images to both groups. Recognition of live images would require greater experience and more rapid appreciation of mucosal barrier defects.

## Conclusion

In conclusion, expert-led didactic learning produces higher accuracy and sensitivity than self-directed audiovisual training. Future eCLE training programs may wish to incorporate didactic teaching into their curriculum despite it being less accessible and more expensive than self-directed learning.

## Competing interests

None

## References

- [1] Chauhan SS, Abu DB, Bhat YM et al. Confocal laser endomicroscopy. *Gastrointest Endosc* 2014; 80: 928–938
- [2] Nguyen NQ, Leong RW. Current application of confocal endomicroscopy in gastrointestinal disorders. *J Gastroenterol Hepatol* 2008; 23: 1483–1491
- [3] Kiesslich R, Goetz M, Vieth M et al. Confocal laser endomicroscopy. *Gastrointest Endosc Clin N Am* 2005; 15: 715–731
- [4] Chang J, Ip M, Yang M et al. The learning curve, interobserver, and intraobserver agreement of endoscopic confocal laser endomicroscopy in the assessment of mucosal barrier defects. *Gastrointest Endosc* 2016; 83: 785–791
- [5] Kiesslich R, Anagnostopoulos GK, Axon A et al. Interobserver variation and standardized training for confocal laser endomicroscopy image interpretation in the upper and lower GI tract. *Gastrointest Endosc* 2007; 5: AB354
- [6] Bertani H, Konda VJ, Arsenescu R et al. Definition of a standardized program of training and credentialing for the use of confocal laser endomicroscopy in gastrointestinal applications: a consensus report. *Gastrointest Endosc* 2014; 79: AB301
- [7] Liu J, Li M, Li Z et al. Learning curve and interobserver agreement of confocal laser endomicroscopy for detecting precancerous or early-stage esophageal squamous cancer. *PLoS One* 2014; 9: e99089
- [8] Tofteland N, Singh M, Gaddam S et al. Evaluation of the updated confocal laser endomicroscopy criteria for Barrett's esophagus among gastrointestinal pathologists. *Dis Esophagus* 2014; 27: 623–629
- [9] Kuiper T, Kiesslich R, Ponsioen C et al. The learning curve, accuracy, and interobserver agreement of endoscope-based confocal laser endomicroscopy for the differentiation of colorectal lesions. *Gastrointest Endosc* 2012; 75: 1211–1217
- [10] Lim LG, Yeoh KG, Salto-Tellez M et al. Experienced versus inexperienced confocal endoscopists in the diagnosis of gastric adenocarcinoma and intestinal metaplasia on confocal images. *Gastrointest Endosc* 2011; 73: 1141–1147
- [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; 106: 1961–1969
- [12] Zhang D, Zhou L, Briggs RO et al. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Inform Manage* 2006; 43: 15–27
- [13] Buchner AM, Gomez V, Heckman MG et al. The learning curve of in vivo probe-based confocal laser endomicroscopy for prediction of colorectal neoplasia. *Gastrointest Endosc* 2011; 73: 556–560
- [14] Rzouq F, Vennalaganti P, Pakseresht K et al. In-class didactic versus self-directed teaching of the probe-based confocal laser endomicroscopy (pCLE) criteria for Barrett's esophagus. *Endoscopy* 2016; 48: 123–127
- [15] Wallace MB, Sharma P, Lightdale C et al. Preliminary accuracy and interobserver agreement for the detection of intraepithelial neoplasia in Barrett's esophagus with probe-based confocal laser endomicroscopy. *Gastrointest Endosc* 2010; 72: 19–24
- [16] Neumann H, Vieth M, Atreya R et al. Prospective evaluation of the learning curve of confocal laser endomicroscopy in patients with IBD. *Histol Histopathol* 2011; 26: 867–872
- [17] Kitchen M. Facilitating small groups: how to encourage student learning. *Clin Teach* 2012; 9: 3–8
- [18] Poole G. The culturally sculpted self in self-directed learning. *Med Educ* 2012; 46: 735–737
- [19] Ozuah PO. Undergraduate medical education: thoughts on future challenges. *BMC Med Educ* 2002; 2: 8
- [20] Roshetsky LM, Coltri A, Flores A et al. No time for teaching? Inpatient attending physicians' workload and teaching before and after the implementation of the 2003 duty hours regulations. *Acad Med* 2013; 88: 1293–1298
- [21] Leong RW, Arshi M, Chang J et al. Acute tumor necrosis factor administration induces transient epithelial cell shredding, mucosal leak and neutrophil recruitment to the gut in a mouse model. *J Gastroenterol Hepatol* 2014; 29: 11
- [22] Leong RW, Wong B, Chen J et al. Intestinal mucosal leakage is detected using in vivo confocal endomicroscopy in macroscopically-normal Crohn's disease and ulcerative colitis. *J Gastroenterol Hepatol* 2012; 27: 109
- [23] Kiesslich R, Duckworth CA, Moussata D et al. Local barrier dysfunction identified by confocal laser endomicroscopy predicts relapse in inflammatory bowel disease. *Gut* 2012; 61: 1146–1153
- [24] Paramosothy S, Leong RWL. Endoscopy: Fluorescein contrast in confocal laser endomicroscopy. *Nat Rev Gastroenterol Hepatol* 2010; 7: 366–368
- [25] Lim LG, Neumann J, Hansen T et al. Confocal endomicroscopy identifies loss of local barrier function in the duodenum of patients with Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis* 2014; 20: 892–900
- [26] Fritscher-Ravens A, Schuppan D, Ellrichmann M et al. Confocal endomicroscopy shows food-associated changes in the intestinal mucosa of patients with irritable bowel syndrome. *Gastroenterology* 2014; 147: 1012–1020
- [27] Chapman CG, Konda VJ. Confocal laser endomicroscopy in inflammatory bowel disease: achieving new depths in mucosal healing. *Gastrointest Endosc* 2016; 83: 792–794
- [28] Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. *Acad Med* 2006; 81: 207–212
- [29] Chu LF, Chan BK. Evolution of web site design: implications for medical education on the internet. *Comput Biol Med* 1998; 28: 459–472
- [30] McLaughlin JE, Roth MT, Glatt DM et al. The flipped classroom: a course redesign to foster learning and engagement in a health professions school. *Acad Med* 2014; 89: 236–243
- [31] Hartsell T, Yuen SC-Y. Video streaming in online learning. *AACE Journal* 2006; 14: 31–43
- [32] Raffellini C. An evaluation of the training activities provided by videoconferencing in a university hospital centre. *Innov Educ Teach Int* 2006; 43: 261–277
- [33] Vaughn L, Baker R. Teaching in the medical setting: balancing teaching styles, learning styles and teaching methods. *Med Teach* 2001; 23: 610–612
- [34] Brown GA, Bull J, Pendlebury M. Assessing student learning in higher education. London: Routledge; 1997
- [35] Hattie J, Timperley H. The Power of Feedback. *Rev Educ Res* 2007; 77: 81–112