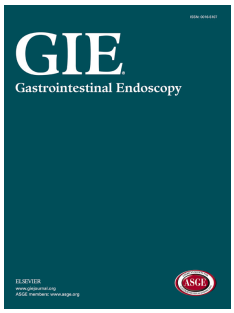


Development and validation of the international blue-light imaging for Barrett's neoplasia classification

Sharmila Subramaniam, MBBS, Kesavan Kandiah, MBChB, Erik Schoon, MD, PhD, Patrick Aepli, MD, Bu' Hayee, MBBS, PhD, Andreas Pischel, MD, Milan Stefanovic, MD, Asma Alkandari, MD, Emmanuel Coron, MD, PhD, Masami Omae, MD, Francisco Baldaque-Silva, MD, PhD, Roberta Maselli, MD, PhD, Raf Bisschops, MD, PhD, Prateek Sharma, MD, PhD, Alessandro Repici, MD, PhD, Pradeep Bhandari, MD, FRCP

PII: S0016-5107(19)32302-8

DOI: https://doi.org/10.1016/j.gie.2019.09.035

Reference: YMGE 11776

To appear in: *Gastrointestinal Endoscopy*

Received Date: 6 June 2019

Accepted Date: 21 September 2019

Please cite this article as: Subramaniam S, Kandiah K, Schoon E, Aepli P, Hayee B', Pischel A, Stefanovic M, Alkandari A, Coron E, Omae M, Baldaque-Silva F, Maselli R, Bisschops R, Sharma P, Repici A, Bhandari P, Development and validation of the international blue-light imaging for Barrett's neoplasia classification, *Gastrointestinal Endoscopy* (2019), doi: https://doi.org/10.1016/j.gie.2019.09.035.

Title:
Development and validation of the international blue-light imaging for Barrett's neoplasia classification

Authors and Affiliations:

1. Sharmila Subramaniam (MBBS)
Department of Gastroenterology, Queen Alexandra Hospital, Portsmouth, UK

2. Kesavan Kandiah (MBChB)
Department of Gastroenterology, Queen Alexandra Hospital, Portsmouth, UK

3. Erik Schoon (MD, PhD)
Department of Gastroenterology, Catharina Hospital, Eindhoven, Netherlands

4. Patrick Aepli (MD)
Department of Gastroenterology & Hepatology, Luzerner Kantonsspital, Luzerne, Switzerland

5. Bu' Hayee (MBBS, PhD)
Department of Gastroenterology, King's College Hospital NHS Foundation Trust, London, UK

6. Andreas Pischel (MD)
Department of Gastroenterology, Sahlgrenska University Hospital, Gothenburg, Sweden

7. Milan Stefanovic (MD)
Department of Gastroenterology, DC Bled, Ljubljana, Slovenia

8. Asma Alkandari (MD)
Department of Gastroenterology & Hepatology, Aljahra Hospital, Kuwait

9. Emmanuel Coron (MD, PhD)
Centre Hospitalier Universitaire & Faculté de Médecine de Nantes
Institut des Maladies de l'Appareil Digestif, France

10. Masami Omae (MD)
Centre for Digestive Diseases, Karolinska University Hospital, Stockholm, Sweden

11. Francisco Baldaque-Silva (MD, PhD)
Centre for Digestive Diseases, Karolinska University Hospital, Stockholm, Sweden

12. Roberta Maselli (MD, PhD)
Digestive Endoscopy Unit, Humanitas Research Hospital, Milan, Italy

13. Raf Bisschops (MD, PhD)
Department of Gastroenterology & Hepatology, Universitaire Ziekenhuizen Leuven, Belgium

14. Prateek Sharma (MD, PhD)
Department of Gastroenterology & Hepatology, Kansas University Medical Center, Kansas, USA

15. Alessandro Repici (MD, PhD)
Digestive Endoscopy Unit, Humanitas Research Hospital, Milan, Italy

16. Pradeep Bhandari (MD, FRCP)
Department of Gastroenterology, Queen Alexandra Hospital, Portsmouth, UK

Corresponding Author

Professor Pradeep Bhandari (MD, FRCP)
Department of Gastroenterology,
Queen Alexandra Hospital
Southwick Hill Road, Cosham
Portsmouth PO6 3LY
United Kingdom

Email: pradeep.bhandari@porthosp.nhs.uk or deep3570@yahoo.co.uk

Title:

Development and validation of the international blue-light imaging for Barrett's neoplasia classification

Abstract

Background and Aims: Detecting subtle Barrett's neoplasia during surveillance endoscopy can be challenging. Blue-light imaging (BLI) is a novel advanced endoscopic technology with high intensity contrast imaging which may improve the identification of Barrett's neoplasia. The aim of this study was to develop and validate the first classification to enable characterisation of neoplastic and non-neoplastic Barrett's using BLI.

Methods: In phase 1, descriptors pertaining to neoplastic and non-neoplastic Barrett's were identified to form the classification (BLINC). Phase 2 involved validation of these component criteria by 10 expert endoscopists assessing 50 BLI images. In phase 3, a web-based training module was developed to enable 15 general (nonexpert) endoscopists to use BLINC. They then validated the classification with an image assessment exercise in phase 4 and their pre- and post-training results were compared.

Results: In Phase 1, the descriptors were grouped into color, pit, and vessel pattern categories to form the classification. In Phase 2, the sensitivity of neoplasia identification was 96.0% with a very good level of agreement among the experts (K=0.83). In Phase 3, 15 general endoscopists completed the training module. In Phase 4, their pretraining sensitivity (85.3%) improved significantly to 95.7% post-training with a good level of agreement (K=0.67).

Conclusion: We developed and validated a new classification system (BLINC) for the optical diagnosis of Barrett's neoplasia using BLI. Despite the limitations of this image-based study with a high prevalence of neoplasia, we believe it has the potential to improve the optical diagnosis of Barrett's neoplasia given the high degree of sensitivity (96%) noted. It is also a promising tool for training in Barrett's optical diagnosis using BLI.

Introduction

Barrett's esophagus (BE) is a precursor to esophageal adenocarcinoma and endoscopic surveillance of BE is now recommended for early detection of neoplasia.[1,2] This enables the use of minimally invasive curative endoscopic techniques that are now first line for management of early Barrett's neoplasia.[3–5]

The current recommended surveillance technique is the Seattle biopsy protocol in which nontargeted quadrantic biopsy specimens are taken every 1 to 2 cm throughout the Barrett's segment.[6] However, this technique has its drawbacks because the biopsies only sample less than 5% of BE, and focal dysplasia can be missed.[7] Adherence to this protocol is poor particularly in long segments of Barrett's, and the vast quantities of biopsy specimens

taken increase cost, prolong procedure time, and reduce patient tolerance to the procedure.[8,9]

The advent of push-button advanced endoscopic imaging technologies over the past decade has provided a means of targeting biopsy samples in order to improve the identification of dysplasia during surveillance endoscopy.[10,11] Narrow-band imaging (NBI; Olympus, Tokyo Japan), I-SCAN (Pentax, Tokyo, Japan) and flexible spectral imaging color enhancement (FICE; Fujifilm, Tokyo, Japan) can enhance mucosal surface and vascular microstructures to facilitate optical diagnosis. In order to determine whether a new technology can indeed replace the current random sampling surveillance protocol, the American Society for Gastrointestinal Endoscopy (ASGE) set out quality thresholds through its Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI) initiative.[12] These PIVI criteria stipulate that a technology needs to achieve a per-patient sensitivity of >90%, specificity >80%, and negative predictive value (NPV) of >98% for dysplasia detection before it can replace quadrantic biopsies.

An ASGE-led meta-analysis showed that acetic acid chromoendoscopy, NBI and confocal laser endomicroscopy (CLE) met the PIVI thresholds set (pooled sensitivity of 96.6% for acetic acid, 94.2% for NBI and 90.4% for CLE).[13] NBI is the most widely studied of these technologies and a recent randomised controlled trial comparing high definition white light to NBI showed that NBI detected a higher proportion of dysplastic areas (30% vs 21%).[14]

Recently, a new optical imaging technology known as Blue Light Imaging (BLI; Fujifilm) was developed. The system uses a 4-light emitting diode multilight technology rather than xenon or laser light sources, thereby producing brighter images. It is the first nonfilter technology producing blue light in a narrow spectrum that is bright enough to identify subtle changes in surface and vessel patterns, which is of relevance in early Barrett's neoplasia. However, to date, no validated classification system exists to guide our use of BLI in the diagnosis of Barrett's neoplasia.

The aim of our study was to develop and validate a classification using BLI for identification of neoplasia in BE.

Methods

Study design

This was a prospective noninterventional image-based study conducted in 4 phases to develop and validate a classification system for Barrett's neoplasia using BLI. Images for this study were collected as part of a trial on Barrett's patients under endoscopic follow-up with institutional board approval (16/ES/0074). The study was carried out in accordance with the Helsinki Declaration.

Acquisition of images for the Blue-Light Imaging Library

The images for this study were captured from 100 patients undergoing Barrett's evaluation procedures in Queen Alexandra Hospital, Portsmouth, United Kingdom. Patients provided written informed consent for these procedures including the use of image and video

recording. The procedures were carried out using the Fujifilm gastroscope series ELUXEO TM 7000 with BLI enhancement (ELUXEO, VP-7000, BL-7000; Fujifilm, Tokyo, Japan) and in each case, the Barrett's segment was examined with high-definition white light before BLI. All patients received a preprocedure drink containing N-acetylcysteine and simethicone to improve mucosal visibility. Images were captured of both neoplastic and non-neoplastic Barrett's and were matched with targeted biopsies from this area for histopathological confirmation. Two independent expert gastrointestinal histopathologists who were not involved in the study and who were blinded to the endoscopic images reported the biopsies of all neoplastic areas. The histopathology report was considered the criterion standard for interpretation in this study with consensus achieved among both histopathologists. Neoplastic Barrett's encompassed areas containing high-grade dysplasia, intramucosal cancer, and submucosally invasive cancer. Non-neoplastic Barrett's had intestinal metaplasia with no dysplasia. No images of low-grade dysplasia (LGD) were included in this study due to the high interobserver variability in histopathological diagnosis and lack of a universal histological definition because it remains an area of controversy. All images were anonymized and stored in a tagged image file format on a secure password protected computer. One hundred high-quality nonmagnification images were selected for inclusion in the study phases (50 non-neoplastic and 50 neoplastic images). All neoplastic visible lesions selected were flat (Paris IIa or IIb)[15] with no nodular components. Images were selected according to their sharpness and clarity with clear visibility of mucosal and vessel patterns. Images from patients with active esophagitis, esophageal strictures, or previous esophageal therapy (endoscopic resection, ablation or radiotherapy) were excluded.

Study Phases

Study phases are shown in Figure 1.

Phase 1: Development of the BLI classification using descriptors pertaining to neoplastic and non-neoplastic Barrett's

Three experienced endoscopists in a tertiary referral center (performed over 200 BE procedures using BLI) reviewed 40 images (20 neoplastic, 20 non-neoplastic) from the library. They proposed a series of descriptors relating to the color, pit, and vessel pattern of neoplastic and non-neoplastic BE. Color was chosen as a category given the brightness of BLI that enabled areas of focal darkness representing neoplasia to be highlighted. Pit and vessel patterns were described according to the type (morphology), distribution (arrangement), and density. After several rounds of voting, complete consensus was achieved among the endoscopists on the choice of descriptors used in these categories to form a structured classification.

Phase 2: Expert validation of proposed descriptors and classification

Ten expert endoscopists (excluding endoscopists in Phase 1) experienced in BE surveillance and neoplasia detection using optical imaging took part in this phase. All had used BLI in their practice. They applied the descriptors to 50 images (25 neoplastic, 25 non-neoplastic) to make a diagnosis. A level of confidence (high or low) was assigned to the predicted diagnosis. All the images were arranged randomly on an online portal developed specifically for this study. They were able to view the descriptors alongside each image and use drop down boxes to select the chosen descriptors. None of the images in Phase 2 were used in Phase 1. The sensitivity, specificity, area under the curve (AUC) and Kappa scores for each

descriptor were calculated. Descriptors with a minimum Kappa value of 0.5 were retained in the classification.

Phase 3: Training Module Development

After the development of the classification, a web-based training module was designed using 10 images of neoplastic and non-neoplastic BE. Each image was illustrated with BLINC descriptors to teach the user how to recognise patterns of neoplastic and non-neoplastic BE on BLI. The various categories (color, pits and vessels) of the classification were demonstrated separately and in combination. The module also contained information on the morphology, location and features of Barrett's neoplasia. The images used in the training module were not included in the assessments.

Phase 4: External validation of the classification by nonexpert endoscopists

Pretraining assessment: 15 endoscopists with exposure to standard BE surveillance in a general population but no experience of optical imaging in BE participated in this phase. They assessed the image library used in Phase 2 and made an optical diagnosis (non-neoplastic versus neoplastic Barrett's) alongside a level of confidence (high/low) for each prediction. They were blinded to the proportion of neoplasia in the exercise and were not given feedback on their diagnostic performance after this phase.

Post-training assessment: The endoscopists completed the BLINC training module and repeated the assessment of the same 50 images (25 neoplastic; 25 non-neoplastic) with an average time period of 8 weeks between pre and post-training tests. The pre- and post-training accuracy, sensitivity, and specificity of neoplasia detection was measured. The proportion of high confidence predictions was also compared between the pre- and post-training assessments.

Statistical Analysis

The sample size was calculated based on the number of observations required for Phase 4. Previous classification studies have demonstrated sensitivity for dysplasia diagnosis at around 80%.[16,17] These studies, like ours, were conducted in tertiary centers with an enriched population and higher prevalence of dysplasia. This study was powered on the basis that the classification would improve the sensitivity of neoplasia diagnosis from 80% to 90%. The data were paired as each assessor evaluated the same images with and without the new classification. It was estimated that 20% of observations would have discordant responses on the 2 timepoints. Using a 5% significance level and 80% power, 155 observations were required for the sensitivity analysis. As an equal number of neoplastic and non-neoplastic images were used, the total number of observations required for the study was doubled (155x2=310). Phase 4 generated 750 observations (15 endoscopists rating 50 images) which satisfied the power calculations.

All data were collected on Microsoft Excel (Microsoft Corporation, Redmond, Wash, USA). Stata version 15.1 (StataCorp, College Station, Tex, USA) was used for statistical analysis. Diagnostic performance was evaluated using sensitivity, specificity, and accuracy. The comparisons between the pre- and post-training data were performed using a bootstrapping approach. Five hundred bootstrap samples were analyzed. The results were used to calculate a confidence interval for each statistic at each timepoint, the difference

between the pre and post-training results, and the corresponding statistical significance. The *P* values were based on a calculation of the standard deviation of the bootstrap samples representing the standard error (SE) of the difference in proportions between timepoints. In turn, a z-test using the observed difference and the SE was used to obtain a *P* value for difference. Interobserver agreement was assessed using the kappa statistic (poor <0.20; fair 0.21-0.40; moderate 0.41-0.60; good 0.61-0.80; very good 0.81-1.00). AUC values of >0.9 were excellent, 0.8-0.9 good, 0.7-0.8 fair, 0.6-0.7 poor, and <0.6 as having no discrimination.

Results

Phase 1: Development of the BLI classification using descriptors pertaining to neoplastic and non-neoplastic Barrett's

A consensus was achieved around 3 main domains: color, pits, and vessels. Each domain had 2 to 3 descriptors as shown in Figure 2. Neoplasia was characterized by the presence of focal darkness, amorphous pits with irregular distribution or pit crowding (increased density) or containing dilated vessels with a noncryptal alignment (Table 1). An image needed to contain an abnormality in at least one domain (color change or distortion in mucosal pits or vessels) to be classified as neoplastic.

Phase 2: Expert validation of proposed descriptors and classification

A total of 500 observations were analyzed (10 endoscopists each assessed 50 images). When BLINC was used, the overall sensitivity, specificity, and accuracy of neoplasia identification was 96%, 94.4% and 95.2%, respectively (Table 2).

The diagnostic performance of each subcategory within the classification was also assessed (Table 3, Figure 4). All subcategories were useful for optical diagnosis of neoplasia with excellent AUC results (>0.90) in the majority of categories (pit type, pit and vessel distribution, density) and good (>0.80) for color and vessel type. Irregular pit distribution showed the highest sensitivity (97.20; 95% CI, 94.32-98.87) and accuracy (95.40; 95% CI, 93.18-97.06) for neoplasia identification. Pit and vessel density (increased = neoplasia) as well as vessel distribution (noncryptal denoting neoplasia) also showed high levels of sensitivity and accuracy exceeding 90%.

The level of agreement between the observers in this phase for each subcategory and the overall diagnosis is displayed in Table 4. The highest level of agreement ("very good," Kappa >0.80) was noted in the scoring of pit distribution and density. Pit type, vessel type, distribution, and density showed a good level of agreement (K=0.61-0.80). Color was the only category where a moderate level of agreement was found (K = 0.59). Nevertheless, when using all categories in the classification, the level of agreement on the overall diagnosis was very good (K = 0.83).

Phase 3: Training Module Development

Figure 3 illustrates aspects of the online training module that the participants undertook. The neoplastic descriptors and importance of each descriptor on its own and in combination was highlighted.

Phase 4: External validation of the classification by nonexpert endoscopists

Seven hundred fifty observations were analyzed in this phase (15 endoscopists assessing 50 images). Their pre- and post-training results were compared (Table 5). There was a significantly higher proportion of high confidence predictions post-training (611/750 or 81.5%) compared with the baseline before BLINC training (441/750 or 58.8%, p<0.0001, Table 6). Sensitivity of neoplasia diagnosis improved significantly post-training reaching levels of 95.7% (p<0.001). However, the specificity of diagnosis decreased from 88.3% to 80.8% (p=0.006) after the use of BLINC. The diagnostic performance of each of the subcategories was analyzed (Table 7) showing that pit type, pit distribution, and vessel distribution displayed the highest specificity for neoplasia (95.2%, 87.2%, and 89.9%, respectively). AUC was excellent for pit distribution and density (0.90), and good for all other subcategories. The diagnostic agreement of neoplasia diagnosis among all endoscopists pretraining was moderate, 0.60 (95% CI, 0.573-0.627). This improved to a good level of agreement post training (K=0.67; 95% CI, 0.646-0.700) although this improvement was not statistically significant (p=0.20). An analysis of interobserver agreement according to each category within the classification was also carried out (Table 8), demonstrating the highest Kappa scores in the pit distribution category (0.74).

When the post-training results were considered alone and stratified according to level of confidence, there was a distinct increase in all diagnostic parameters (sensitivity 96%, specificity 86%) with a much higher proportion of neoplastic images in the high confidence cohort (Table 9).

Discussion

We have proposed and validated the first classification system using BLI in the identification of Barrett's neoplasia. A high level of sensitivity (96%) for neoplasia was reached when validated among expert endoscopists. This was matched by nonexpert endoscopists after training where 96% sensitivity for neoplasia diagnosis was also achieved. This, coupled with high interobserver agreement (K>0.8 in our study) further strengthens the validity and reproducibility of BLINC across a broad user spectrum.

The use of optical imaging in BE is not new. The most widely researched imaging technology so far in this field is NBI and several iterations of validated classifications have been proposed over the past decade.[18,19,20] However, subsequent validation of some of these classifications have shown only modest results, perhaps in part due to their complexity limiting uptake in the community.[21] Recently, a new simplified consensus driven NBI classification for Barrett's (BING) was developed and validated by an international working group of experts.[17] The BING criteria identified patients with dysplasia with an overall 85% accuracy, 80% sensitivity and 88% specificity. Higher levels of accuracy, sensitivity, and specificity (92%, 91%, 93%, respectively) were demonstrated when this classification was used in high confidence predictions with good interobserver agreement (K=0.68). However, when BING was used with different modalities (I-scan, magnification, and/or acetic acid) accuracy rates were lower (accuracy of I-scan with acetic acid 79%, magnification and acetic acid 83%).[16] Therefore, we did not incorporate classifications designed for NBI or i-scan in the design of this study to assess the utility of BLI in identifying Barrett's neoplasia.

The mode of action of BLI differs from that of existing technologies as it uses light emitting diodes to directly emit a blue light without involving a narrow-band filter or digital postprocessing technology. Given the differences in this technology it was important that a bespoke classification was designed to facilitate interpretation of BLI images and for training. The contrast-enhancing properties and brightness of the BLI image of BLI enable it to be used to recognize areas of dysplasia within BE. A recent study showed that experts using BLI were able to improve their performance in delineating neoplastic lesions compared with white-light endoscopy.[22] Our classification is unique in its inclusion of "color" as a category in addition to pit and vessel patterns. We describe "focal darkness" associated with neoplasia as we believe it can be used as a red flag that draws the endoscopist's attention to an abnormal area, which is then categorized further by identifying the pit and vessel structure. Focal darkness was correctly associated with neoplasia in 84.8% of neoplastic images rated by the 10 experts in our study with moderate interobserver agreement (K=0.59). Another unique feature of BLINC is its incorporation of the subcategories of type, distribution, and density thereby providing a framework to analyse BE mucosa and enable the endoscopist to distinguish between neoplastic and non-neoplastic patterns. This is particularly important when implementing a new technology among nonexperts without the experience or intuition built upon years of experience to confidently distinguish between a neoplastic and non-neoplastic pit and vascular structure. The expert validation of this classification showed that nonuniform or irregular pits, increased pit density, "noncryptal" vessels and increased vessel density were the best predictors of neoplasia (sensitivities of 97%, 94%, 92%, and 91%, respectively). These parameters also displayed higher levels of interobserver agreement (K >0.8 for pit distribution and density, K>0.7 for vessel distribution and density).

One of the strengths of this study is the incorporation of an external validation phase (Phase 4) to assess applicability of the classification among general (nonexpert) endoscopists with no specific experience of Barrett's neoplasia detection and no routine use of optical imaging in their daily practice. Our results showed that with adequate training this group was also able to reliably identify (sensitivity=95.7%) and rule out neoplasia in the images assessed. The decrease in specificity in this phase from 88.3% to 80.8% after implementation of BLINC is likely due to the increased awareness and focus on neoplastic features leading to a tendency to overcall neoplasia. However, this is still preferable to the baseline of this group where a higher neoplasia miss rate was found. We also demonstrated that the use of BLINC significantly increased the proportion of high confidence responses (81.5%) among general endoscopists, almost reaching the level noted in the expert group (88.4%). Even in low confidence predictions made in Phase 4, high levels of sensitivity were shown (94%), suggesting that BLINC has a role to play in improving detection and reliably excluding neoplasia. Although these results are encouraging and come close to the pre-defined ASGE PIVI thresholds, further prospective in-vivo studies in a surveillance population will need to be carried out for optimal validation. At present, the results of this study are not generalizable to the general surveillance population with a much lower prevalence of neoplasia and at present, BLI may not be used to replace standard surveillance protocol biopsies.

It is also important to note that these results were achieved with the application of BLINC through a web-based training portal rather than with face-to-face teaching sessions. The

absence of immediate direct feedback with no opportunity for discussion may have underused the strengths of the classification, and it is possible that further improvement in diagnostic performance could have been reached with a different training method. Nevertheless, it is encouraging to note the strength of results achieved via web learning, which may favor the dissemination of this technology and classification for use in daily practice outside expert centers.

Another strength of this classification is the overall level of agreement, which was good (K=0.67) in nonexperts and very good in experts (K=0.83). This compares favorably with other classifications including BING (K=0.68) in which validation was conducted with experts only. A study by Silva et al[23] demonstrated that the overall interobserver agreement among endoscopists using the Kansas, Amsterdam, and Nottingham NBI classifications was 0.44, 0.47 (moderate), and 0.34 (fair). We used images rather than videos in all phases of this study. Although videos may simulate real-time procedures better, the objective of this study was to establish whether the proposed BLI classification could be used reliably to predict neoplasia, and a high-quality still image served better for this purpose. Furthermore, even in a real-life scenario using BLI, the endoscopist would capture and freeze an image of any potentially neoplastic areas in order to interrogate its pit and vessel pattern to aid in making an optical diagnosis.

This study has several limitations. First, we did not incorporate a real-time in-vivo validation phase in a Barrett's surveillance population. However, the intention of this study was to develop a new classification for a novel technology, so it would have been premature to include a real-time in vivo validation phase at this stage. Second, an equal proportion of neoplastic and non-neoplastic images were used in this study; such an enriched concentration of neoplasia does not reflect the reality of surveillance in a community-based setting, and therefore negative and positive predictive values were not reported. It is important to highlight that the main purpose of this study was to develop a structured method of recognizing neoplasia and therefore a higher prevalence was required, which would not have been found in a surveillance population. Participants were not aware of the proportion of neoplasia and despite the enriched sample, the high sensitivity obtained is a strong indicator of the classification's performance. Nevertheless, further validation of BLINC will be required in a BE surveillance population. Third, we did not include images of LGD. LGD remains a challenge for endoscopists and pathologists due to the high interobserver variability.[24] The histopathological criteria for distinguishing LGD from inflammation are not well established and to date, no clear endoscopic classification of LGD exists, which is why we did not include images of LGD in this study. Robust universal histological definitions of LGD need to be established in order to be able to interpret endoscopic appearances of LGD, particularly because histological confirmation remains the criterion standard for dysplasia diagnosis and guides decisions on endoscopic therapy, including ablation. As an endoscopic and histopathological description of LGD evolves, it can be incorporated into the classification in the future. We also acknowledge the selection bias created by a limited number of images, expert driven consensus classification, and high quality of images selected, which may have accentuated the positive results achieved. The distinctive categorical nature of the images selected (either neoplastic or non-neoplastic) may not mirror the uncertainties of real-time examination, which could affect the performance of the classification when used in real-time validation. Although the same set

of images were used to validate both pre- and post-training components of Phase 4, we minimized the effect of any recall bias by introducing a gap between pre- and post-training tests, a different random order to the images rated in the post-training phase and blinding the endoscopists to the histology until this phase was completed. Finally, we do acknowledge that because BLI is a new technology that has only recently been approved for use, it is unlikely to be widely available to most endoscopists at this time. Nevertheless, this makes it an ideal time to introduce a classification system to aid future users of this new technology.

In conclusion, we have developed the first classification system (BLINC) using a novel BLI technology in BE. We also designed and implemented an online training tool on BLINC and showed that this new classification can be used effectively by both experts and nonexperts. BLI is a promising tool for optical diagnosis in Barrett's, although further studies need to be undertaken real-time in a general Barrett's surveillance population.

Acknowledgements

The authors would like to acknowledge Dr James Brown and Sarah Cooper from the University of Portsmouth for their facilitation in hosting the training module using the Moodle platform and Paul Bassett for his support with statistical analysis. We would also like to thank the consultants, fellows and nurse endoscopists from UK, Italy and USA who contributed to the validation of this classification.

References

1. Spechler SJ, Sharma P, Souza RF, Inadomi JM, Shaheen NJ. American Gastroenterological Association medical position statement on the management of Barrett's esophagus. Gastroenterology. 2011;140:1084–91.
2. Fitzgerald RC, Di Pietro M, Ragunath K, Ang Y, Kang JY, Watson P, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. Gut. 2014;63:7–42.
3. Weusten B, Bisschops R, Coron E, Dinis-Ribeiro M, Dumonceau J-M, Esteban J-M, et al. Endoscopic management of Barrett's esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy. 2017;49.
4. Sharma P, Katzka DA, Gupta N, Ajani J, Buttar N, Chak A, et al. Quality Indicators for the Management of Barrett's Esophagus, Dysplasia, and Esophageal Adenocarcinoma: International Consensus Recommendations from the American Gastroenterological Association Symposium. In: Gastroenterology. 2015.
5. Phoa KN, Pouw RE, Bisschops R, Pech O, Ragunath K, Weusten BL a M, et al. Multimodality endoscopic eradication for neoplastic Barrett oesophagus: results of an European multicentre study (EURO-II). Gut. 2015;
6. Levine DS, Blount PL, Rudolph RE, Reid BJ. Safety of a systematic endoscopic biopsy protocol in patients with Barrett's esophagus. Am J Gastroenterol. 2000;95:1152–7.
7. Kariv R, Plesec TP, Goldblum JR, Bronner M, Oldenburgh M, Rice TW, 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;7:653–8.

8. Abrams JA, Kapel RC, Lindberg GM, Saboorian MH, Genta RM, Neugut AI, et al. Adherence to Biopsy Guidelines for Barrett's Esophagus Surveillance in the Community Setting in the United States. Clin Gastroenterol Hepatol. 2009;7:736–42.
9. Westerveld D, Khullar V, Mramba L, Ayoub F, Brar T, Agarwal M, et al. Adherence to quality indicators and surveillance guidelines in the management of Barrett's esophagus: a retrospective analysis. Endosc Int Open [Internet]. 2018;06(03):E300–7. Available at: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0044-101351
10. Subramanian V, Ragunath K, Hirchowitz BI, Peters CW, Curtiss LE, Kwon RS, et al. Advanced endoscopic imaging: a review of commercially available technologies. Clin Gastroenterol Hepatol [Internet]. 2014;12(3):368–76.e1. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23811245
11. Qumseya BJ, Wang H, Badie N, Uzomba RN, Parasa S, White DL, et al. Advanced imaging technologies increase detection of dysplasia and neoplasia in patients with barrett's esophagus: A meta-analysis and systematic review. Clinical Gastroenterology and Hepatology. 2013.
12. Abu Dayyeh BK, Thosani N, Konda V, Wallace MB, Rex DK, Chauhan SS, et al. ASGE technology committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. Vol. 81, Gastrointestinal Endoscopy. 2015. p. 502–16.
13. Thosani N, Abu Dayyeh BK, Sharma P, Aslanian HR, Enestvedt BK, Komanduri S, et al. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE Preservation and Incorporation of Valuable Endoscopic Innovations thresholds for adopting real-time imaging-assisted endoscopic targeted biopsy during endoscopic surveillance. Gastrointest Endosc. 2016;83:684–698.e7.
14. Sharma P, Hawes RH, Bansal A, Gupta N, Curvers W, Rastogi A, et al. Standard endoscopy with random biopsies versus narrow band imaging targeted biopsies in Barrett's oesophagus: A prospective, international, randomised controlled trial. Gut. 2013;62:15–21.
15. Lambert R, Lightdale CJ. The Paris endoscopic classification of superficial neoplastic lesions: Esophagus, stomach, and colon - Paris, France November 30 to December 1, 2002: Foreword. In: Gastrointestinal Endoscopy. 2003.
16. Lipman G, Bisschops R, Sehgal V, Ortiz-Fernández-Sordo J, Sweis R, Esteban JM, et al. Systematic assessment with I-SCAN magnification endoscopy and acetic acid improves dysplasia detection in patients with Barrett's esophagus. Endoscopy. 2017;49:1219–28.
17. Sharma P, Bergman JJGHM, Goda K, Kato M, Messmann H, Alsop BR, et al. Development and Validation of a Classification System to Identify High-Grade Dysplasia and Esophageal Adenocarcinoma in Barrett's Esophagus Using Narrow-Band Imaging. Gastroenterology [Internet]. 2016;150(3):591–8. Available at: http://dx.doi.org/10.1053/j.gastro.2015.11.037
18. Singh R, Anagnostopoulos GK, Yao K, Karageorgiou H, Fortun PJ, Shonde A, et al. Narrow-band imaging with magnification in Barrett's esophagus: Validation of a simplified grading system of mucosal morphology patterns against histology. Endoscopy. 2008;40:457–63.
19. Kara MA, Ennahachi M, Fockens P, ten Kate FJW, Bergman JJGHM. Detection and classification of the mucosal and vascular patterns (mucosal morphology) in Barrett's esophagus by using narrow band imaging. Gastrointest Endosc. 2006;

20. Sharma P, Bansal A, Mathur S, Wani S, Cherian R, McGregor D, et al. The utility of a novel narrow band imaging endoscopy system in patients with Barrett's esophagus. Gastrointest Endosc. 2006;64:167–75.
21. Baldaque-Silva F, Marques M, Lunet N, Themudo G, Goda K, Toth E, et al. Endoscopic assessment and grading of Barrett's esophagus using magnification endoscopy and narrow band imaging: Impact of structured learning and experience on the accuracy of the Amsterdam classification system. Scand J Gastroenterol. 2013;48:160–7.
22. de Groof AJ, Swager A-F, Pouw RE, Weusten BLAM, Schoon EJ, Bisschops R, et al. Blue-light imaging has an additional value to white-light endoscopy in visualization of early Barrett's neoplasia: an international multicenter cohort study. Gastrointest Endosc [Internet]. 2019 Jan 9; Available at: https://doi.org/10.1016/j.gie.2018.10.046
23. Silva FB, Dinis-Ribeiro M, Vieth M, Rabenstein T, Goda K, Kiesslich R, et al. Endoscopic assessment and grading of Barrett's esophagus using magnification endoscopy and narrow-band imaging: Accuracy and interobserver agreement of different classification systems (with videos). Gastrointest Endosc. 2011;73:7–14.
24. Vennalaganti P, Kanakadandi V, Goldblum JR, Mathur SC, Patil DT, Offerhaus GJ, et al. Discordance Among Pathologists in the United States and Europe in Diagnosis of Low-Grade Dysplasia for Patients With Barrett's Esophagus. Gastroenterology. 2017;152:564–570.e4.

Figure Legends

Figure 1. Study phases
Figure 2. Blue-light imaging in Barrett's neoplasia classification (BLINC).
Figure 3: Integrated example of BLINC descriptor application to images of neoplastic and non-neoplastic Barrett's.
Figure 4. Receiver operating characteristics for all subcategories validated in Phase 2.

Tables

Table 1: Neoplastic and non-neoplastic descriptors used in BLINC

Domains	Descriptors	Optical diagnosis in BE
Color	Focal darkness	Neoplastic
	No focal darkness	Non-neoplastic
Pits	Amorphous, circular, tubular or gyriform pits, irregular distribution or increased density	Neoplastic
	Circular, tubular or gyriform pits, regular distribution and density	Non-neoplastic
Vessels	Dilated/branching pattern, noncryptal distribution, increased or focal loss of density	Neoplastic
	Non dilated or branching pattern, pericryptal vessels with normal density	Non-neoplastic

Table 2: Diagnostic performance of BLINC when used by experts

	Proportion of neoplasia	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)
Overall	50%	96.00 (92.77-98.07)%	94.40 (90.78-96.90)%	95.20 (92.94-96.90)%
High confidence predictions (n=442, 88.4%)	47.51 (42.77-52.28)%	98.57 (95.88-99.70)%	97.41 (94.46-99.05)%	97.96 (96.17-99.06)%
Low confidence predictions (n=58, 11.6%)	68.97 (55.46-80.46)%	82.50 (67.22-92.66)%	55.56 (30.76-78.47)%	74.14 (60.96-84.74)%

Table 3: Performance of each subcategory within BLINC for the diagnosis of neoplasia by experts

Categories	Subcategories	Area under the curve (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Color	Focal darkness vs no focal darkness	0.86 (0.83-0.90)	84.80 (79.74-89.01)	87.20 (82.41-91.08)
Pits	**Type** (amorphous vs gyriform/ tubular/ circular)	0.93 (0.90-0.96)	89.60 (85.13-93.09)	96.40 (93.28-98.34)
	Distribution (irregular vs regular)	0.95 (0.93-0.98)	97.20 (94.32-98.87)	93.60 (89.81-96.30)
	Density (Increased vs normal)	0.94 (0.92-0.97)	94.80 (91.27-97.20)	94.00 (90.30-96.60)
Vessels	**Type** (dilated vs nondilated)	0.89 (0.85-0.92)	84.40 (79.30-88.67)	92.80 (88.86-95.68)
	Distribution (noncryptal vs pericryptal)	0.93 (0.90-0.95)	92.40 (88.39-95.36)	92.80 (88.86-95.68)
	Density (Increased vs normal)	0.92 (0.89-0.95)	91.20 (86.98-94.40)	92.80 (88.86-95.68)

Table 4: Interobserver variability among 10 experts validating BLINC

Criterion	Kappa value
Color	0.586 (0.544-0.627)
Pit type (amorphous vs nonamorphous)	0.785 (0.743-0.826)
Pit distribution	0.842 (0.801-0.884)
Pit density	0.801 (0.760-0.842)
Vessel type	0.688 (0.646-0.729)
Vessel distribution	0.768 (0.727-0.809)
Vessel density	0.742 (0.701-0.783)
Overall diagnosis	0.831 (0.790-0.872)

Table 5: Pre- and post-BLINC training results in nonexpert endoscopists

	Pretraining	**Post-training**	**Change**	**Significance**
Sensitivity (95% CI)	85.3 (81.1-88.2)%	95.7 (93.3-97.6)%	10.4 (6.1-14.6)	p<0.001
Specificity (95% CI)	88.3 (85.0-91.7)%	80.8 (76.8-85.1)%	-7.5 (-12.4, -2.3)	p=0.006
Accuracy (95% CI)	86.8 (84.0-88.9)%	88.3 (86.1-90.7)%	1.5 (-1.9, 4.9)	p=0.42

Table 6: Proportion of predictions according to level of confidence pre and post BLINC training by nonexpert endoscopists

	Pretraining	**Post-training**	**Change**	**Significance**
High confidence predictions	441/750 (58.8%)	611/750 (81.5%)	22.7 (18.1-27.1)	p<0.0001
Low confidence predictions	309/750 (41.2%)	139/750 (18.5%)	22.7 (18.1-27.1)	p<0.0001

Table 7: Diagnostic performance of the individual descriptor subcategories in predicting neoplasia when used by nonexpert endoscopists

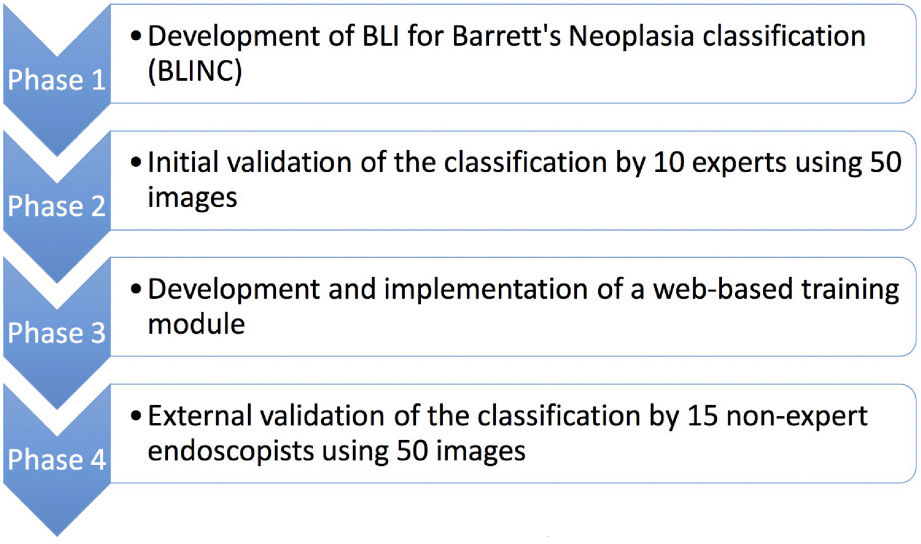
Categories	Subcategories	Area under the curve (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Color	Focal darkness vs no focal darkness	0.82 (0.79-0.86)	82.67 (78.45-86.36)%	82.67 (78.45-86.36)%
Pits	**Type** (amorphous vs gyriform/ tubular/ circular)	0.82 (0.79-0.85)	69.33 (64.39-73.96)%	95.20 (92.52-97.13)%
	Distribution (irregular vs regular)	0.91 (0.88-0.93)	94.40 (91.57-96.50)%	87.20 (83.39-90.41)%
	Density (Increased vs normal)	0.90 (0.87-0.92)	93.60 (90.63-95.86)%	85.87 (81.92-89.23)%
Vessels	**Type** (dilated vs nondilated)	0.84 (0.81-0.87)	83.73 (79.60-87.32)%	84.27 (80.18-87.80)%
	Distribution (noncryptal vs pericryptal)	0.85 (0.82-0.88)	80.80 (76.44-84.66)%	89.87 (86.36-92.73)%
	Density (Increased vs normal)	0.87 (0.84-0.90)	89.87 (86.36-92.73)%	84.53 (80.47-88.04)%

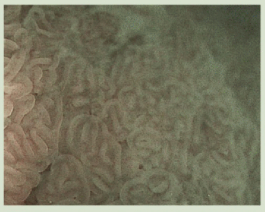
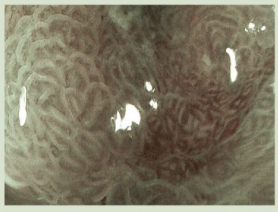
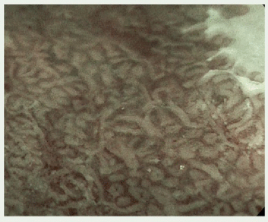
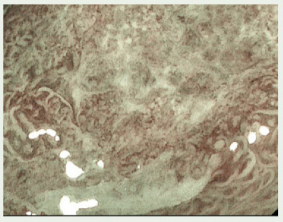
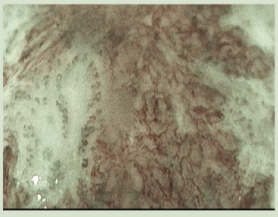
Table 8: Interobserver variability among nonexpert endoscopists

Criterion	Kappa value
Color	0.501 (0.474-0.528)
Pit type (amorphous vs nonamorphous)	0.622 (0.595-0.649)
Pit distribution	0.742 (0.715-0.769)
Pit density	0.621 (0.594-0.648)
Vessel type	0.457 (0.430-0.485)
Vessel distribution	0.643 (0.616-0.670)
Vessel density	0.630 (0.603-0.657)
Overall diagnosis	0.673 (0.646-0.700)

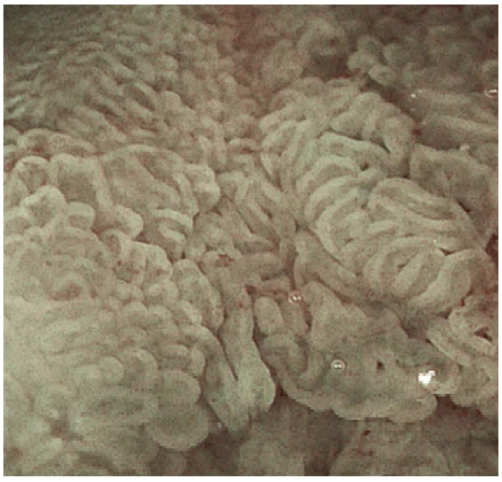
Table 9: Diagnostic performance for neoplasia identification stratified according to level of confidence (low vs high) after adoption of BLINC by nonexpert endoscopists

Confidence level	Proportion of neoplasia	Sensitivity	Specificity	Accuracy
High confidence (n=611)	81.50%	95.94 (93.15-97.82)%	86.25 (81.76-89.99)%	91.33 (88.81-93.43)%
Low confidence (n=139)	40.15%	94.55 (84.88-98.86)%	60.98 (49.57-71.56)%	74.45 (66.30-81.52)%



		Non Neoplastic	Neoplastic
Colour		No focal darkness	Focal darkness
			
Mucosal pits (crypts)	**Type**	Circular, tubular, gyriform	Amorphous, circular, tubular, gyriform
	Distribution	Regular	Irregular
	Density	Normal	Increased
			
Vessels	**Type**	Non dilated, non branching	Dilated, branching
	Distribution	Pericryptal	Non cryptal
	Density	Normal	Increased or focal loss
			

Non Neoplastic Barrett's



Mucosal Pits

- Type: Circular & tubular
- Distribution: Uniform
- Density: Normal

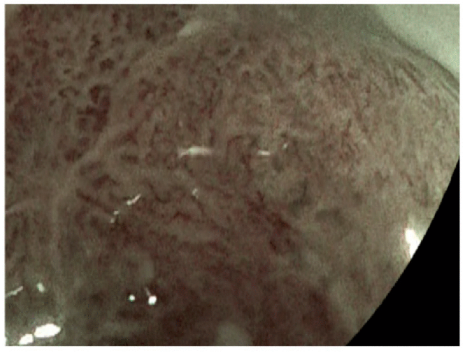
Vessels

- Type: Non dilated
- Distribution: Pericryptal
- Density: Normal

Colour

- No focal darkness

Neoplastic Barrett's



Mucosal Pits

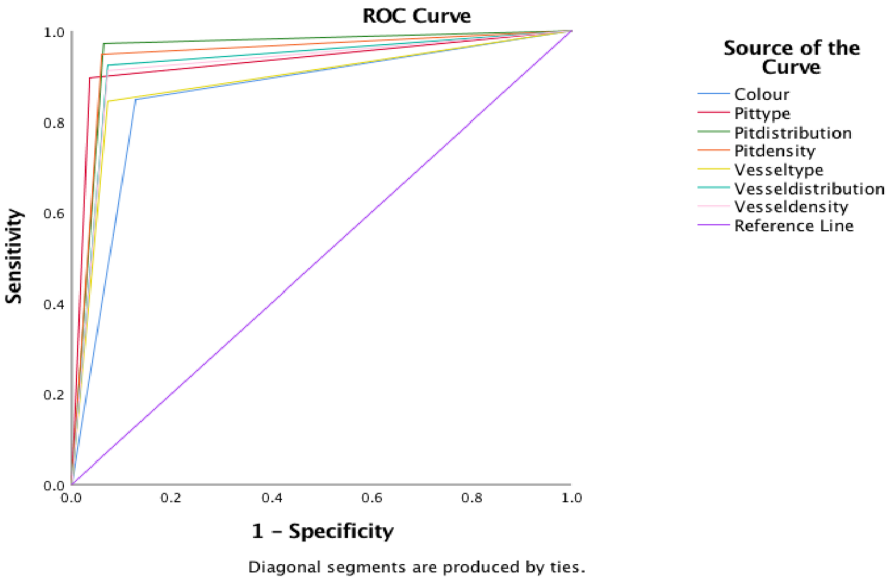
- Type: Amorphous
- Distribution: Focal loss
- Density: Increased

Vessels

- Type: Dilated & tortuous
- Distribution: Non cryptal
- Density: Increased (new vessels)

Colour

- Dark



Acronyms and Abbreviations

ASGE - American Society of Gastrointestinal Endoscopy

AUC - Area under the curve

BE - Barrett's oesophagus

BING - Barrett's International NBI Group

BLI - Blue Light Imaging

BLINC - Blue Light Imaging for Barrett's Neoplasia Classification

CI - Confidence Interval

CLE - Confocal Laser Endomicroscopy

FICE - Flexible spectral Imaging Colour Enhancement

LGD - Low grade dysplasia

NBI - Narrow Band Imaging

NPV - Negative Predictive Value

PPV - Positive Predictive Value

PIVI - Preservation and Incorporation of Valuable Endoscopic Innovations