Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy

Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy

A. Swager[1], F. van der Sommen[2], S.R. Klomp[2], S. Zinger[2], S.L. Meijer[3], E.J. Schoon[4], J.J. Bergman[1], P.H.N. de With[2], W.L. Curvers[4]

1 Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, the Netherlands
3 Department of Pathology, Academic Medical Center, Amsterdam, the Netherlands
2 Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands
4 Department of Gastroenterology and Hepatology, Catharina Hospital, Eindhoven, the Netherlands

---

ABSTRACT

Background and aims: Volumetric laser endomicroscopy (VLE) is an advanced imaging system that provides a near-microscopic resolution scan of the esophageal wall layers up to 3 mm deep. VLE has the potential to improve detection of early neoplasia in Barrett's esophagus (BE). However, interpretation of VLE images is complex due to the large amount of data that needs to be interpreted in real-time. The aim of this study was to investigate the feasibility of a computer algorithm to identify early BE neoplasia on ex vivo VLE images.

Methods: We used 60 VLE images from a database of high-quality ex vivo VLE-histology correlations, obtained from BE patients ± neoplasia (30 non-dysplastic (ND)BE and 30 HGD/EAC images). VLE features from a recently developed clinical VLE prediction score for BE neoplasia served as input for the algorithm: 1) higher VLE surface than subsurface signal, 2) lack of layering. With this input novel clinically inspired algorithm features were developed, based on signal intensity statistics and grayscale correlations. For comparison, generic image analysis methods were examined for their performance to detect neoplasia. For classification of the images in the NDBE or neoplastic group, several machine learning methods were evaluated. Leave-one-out cross-validation was employed for algorithm validation.

Results: Three novel clinically inspired algorithm features were developed. The feature 'layering and signal decay statistics' showed the optimal performance compared to the other clinically features ('layering' and 'signal intensity distribution') and generic image analyses methods, with an area under the receiver operating curve (AUC) of 0.95. Corresponding sensitivity and specificity were 90% and 93%, respectively. In addition, the algorithm showed a better performance than the clinical VLE prediction score (AUC 0.81).

Conclusion: This is the first study in which a computer algorithm for BE neoplasia was developed based on VLE images with direct histological correlates. The algorithm showed good performance to detect BE neoplasia in ex vivo VLE images compared to the performance of a recently developed clinical VLE prediction score. This study suggests that an automatic detection algorithm has the potential to assist endoscopists to detect early neoplasia on VLE. Future studies on in vivo VLE scans are needed to further validate the algorithm.

INTRODUCTION

Esophageal adenocarcinoma (EAC) is among the top ten of most common and lethal cancers in the world.[1] Barrett's esophagus (BE) is a known precursor lesion for EAC and therefore BE patients currently undergo regular surveillance with white light endoscopy and random biopsies.[2,3] However, an unknown number of early neoplastic lesions in the BE segment are missed due to their subtle appearance and sampling error. Volumetric laser endomicroscopy (VLE) is a novel advanced imaging system that incorporates second generation optical coherence tomography (OCT) technology.[4] Improvements in image acquisition speed enable this balloon-based system to perform a quick circumferential scan of the entire distal esophagus.[5] VLE thus provides a three-dimensional map of near-microscopic resolution of the subsurface layers of the esophagus. Therefore, VLE may increase the detection of early neoplasia and thereby improve current BE surveillance protocols.

VLE scans comprise a large amount of data that needs to be interpreted in real-time during endoscopy. Therefore, interpretation of VLE in the endoscopy suite may be complex and time consuming. Recently, clinical VLE prediction models were developed for BE neoplasia showing good accuracy.[6,7] However, pilot studies using ex vivo and in vivo VLE data may suggest that full scan VLE interpretation by experts remains difficult.[8,9] In vivo case reports suggested that VLE could detect additional neoplasia, however, clinical trials in larger sample size are necessary to confirm these findings.[10] We envision that development of quantitative methods that enable red-flagging of areas suspicious for neoplasia, is important to aid endoscopists in real-time interpretation of VLE scans. In this study we examine such a quantitative method focusing on computer-aided detection. Detection algorithms have the potential to analyze all pixels in a VLE scan in a short time period, and machine learning methods can be used to train the algorithm to recognize abnormal patterns associated with neoplasia. Recently, van der Sommen *et al* developed a computer algorithm for the detection of early neoplasia on endoscopic images from BE patients, showing a reasonable accuracy to identify BE neoplasia.[11] Other studies have shown that image analysis and computer-aided diagnosis can aid in classification of neoplasia in OCT images.[12,13] The aim of this study was to investigate the feasibility of a computer algorithm to detect early BE neoplasia on ex vivo VLE images.

METHODS

Patients and image database

The VLE images used in this study were derived from a database of high-quality one-to-one VLE-histology correlations. The construction of this database has been described previously.[14] In short, endoscopic resection specimens from Barrett's patients with and without early neoplasia were used. The fresh specimens were scanned ex-vivo with VLE in a custom-made fixture. Guided by previously placed in- vivo and ex-vivo markers, histological sectioning was performed. Histology slides and VLE frames were considered to be correlated one-to-one when ≥ 2 markers were visible on both modalities. Within the correlated VLE images, areas of interest were selected containing one mucosal type. In total 60 images containing areas of interest of sufficient quality were included: 30 non-dysplastic (ND)BE and 30 neoplastic images, containing high-grade dysplasia (HGD) or early esophageal adenocarcinoma (EAC). In figure 1, two example VLE images of both non-dysplastic (ND)BE and neoplasia are shown.

Nvision VLE™ Imaging System

The VLE system consists of a disposable optical probe, inflation system and imaging console, which consists of a display with user interface, a swept light source, optical receiver and interferometer, and a data-acquisition computer (see figure 2). The images used in this study were rendered from scans made by the pre-commercial and the commercial VLE imaging system. Apart from technological changes and a different outer diameter of the balloon (20 vs 25 mm), there were no significant differences in image quality for analysis. The light source consists of near-infrared light (wave length range 1250 to 1350 nm) that is transmitted into the catheter. A polymeric, noncompliant balloon at the distal end centers the optical probe in the esophageal lumen for in vivo imaging. During an automatic pullback of the optical probe a 6-cm long segment of the esophagus is scanned circumferentially with VLE in 90 seconds, reaching a depth penetration of 3 mm into the tissue. The axial and lateral resolutions are 7 μm and 40 μm, respectively. The NvisionVLE™ Imaging System (NinePoint Medical Inc., Cambridge, MA, USA) incorporates the second generation optical coherence tomography (OCT) technology, termed optical frequency-domain imaging technology. For more comprehensive technical details of, we refer to previous publications.[15–17]

Development of a computer algorithm for BE neoplasia detection on VLE images

General framework

A computer-aided detection algorithm enables analysis of all pixels in a VLE scan and can be trained to recognize abnormal patterns. Figure 3 depicts the general framework for the development of a

VLE computer algorithm, and the different steps are described in the figure legend. Several machine learning methods (see step 3 in figure 3) were applied to train the algorithm to classify VLE images into the neoplastic or the NDBE group, according to different VLE image features. For validation of the algorithm, leave-one-out-cross validation was applied, in which the algorithm was subsequently trained with all images but one and evaluated on that left-out image. This process was repeated for each image independently in order to obtain a score for each image. The advantage of an algorithm over human observers is that it can be reset for each image to retrain it from the beginning.

Evaluation of feature extraction and machine learning methods

In other automated image analysis tasks (e.g. face recognition and traffic sign detection) several generic image feature descriptors and machine learning methods have been very successful. Therefore, several generic image analysis features were evaluated for the task of feature extraction from the VLE images and evaluated for the classification of the VLE images in the NDBE or neoplastic group. All feature extraction and machine learning methods used during the development of the algorithm are listed in table 1. Extensive explanation of these technical methods falls outside the scope of this manuscript. For a comprehensive overview and more technical details of the employed features and methods, we refer to the technical publication by our group.[18]

Clinically inspired algorithm features

Next to evaluating generic methods that are widely used in the field of image analysis, we implemented specific clinical knowledge on VLE features to develop more substantiated algorithm features specifically for VLE data. Previously identified VLE features predictive for BE neoplasia served as clinical input for the algorithm: 1) lack of layering, 2) higher VLE surface signal than subsurface signal in tissue [6]. Figure 1 and 4 depict these features in two VLE images containing NDBE and neoplasia. Inspired by these clinical VLE features, three novel algorithm features specific for VLE were developed. The first novel feature was termed “layering”, and was constructed by calculating a VLE signal intensity histogram using 8 intensity categories over a number of layers (see figure 5). The second feature termed “layering and signal decay statistics”, was constructed based on gray-level pixel statistics. The third feature termed “signal intensity distribution” utilized average intensities of the image. In a technical publication by our group, all technical details of the aforementioned features have been described.[18]

Comparison of clinically inspired features versus generic image analysis features

The algorithm development process consisted of the following steps. First, the performance of different generic image analysis features to distinguish neoplasia on VLE images was compared with the performance of the novel clinically inspired image analysis features. This provided a comprehensive overview of features that were suitable for the classification of VLE images (table 1). Second, the most promising image analysis features were further optimized, for more technical details we refer to the technical publication.[18] All experiments were carried out using software package Matlab 2016a (Mathworks Inc., Natick, Massachusetts, USA) on a desktop PC (hexa-core @3.3 GHz CPU, 16GB RAM, 2GB GPU).

Comparison of computer-aided detection versus VLE experts

Finally, the algorithm performance was compared to the performance of VLE experts. In a previous study by our group, the same images used in this study were scored by VLE experts for the presence of neoplasia, using the clinical VLE features layering, surface signal and presence of irregular glands.[6] Blinded for the corresponding histology, two VLE experts scored 20 images in a learning phase and 40 images in a validation phase. The data from that study was compared with the performance of the computer algorithm for correctly identifying thee VLE images containing BE neoplasia.

Statistical analysis

SPSS 23 Software for Windows was used to perform statistical analysis. Standard descriptive statistics were used to describe patient characteristics and endoscopic findings: mean (± standard deviation (SD) was used in case of a normal distribution of variables. For evaluation of the performance of the algorithm, sensitivity/specificity analysis was applied, whereby the area under the receiver operating characteristic (ROC) curve (AUC) was calculated. By varying the cut-off value for neoplasia, the ROC was computed. The histological diagnosis, correlated to the images served as reference standard.

RESULTS

Patients and VLE-histology database

The previously constructed VLE-histology database consisted of 52 ER specimens from 29 BE patients (mean age 67 years (SD ± 8.4); 22 males). Worst histological diagnosis per patient was NDBE in 6, low-grade dysplasia (LGD) in 2, HGD in 5 and EAC in 16. In total, 86 VLE-histology matches containing 200 areas of interest were constructed. For this study, 30 NDBE and 30 HGD/EAC areas of interest of sufficient quality were selected. These 60 images were derived from 19 different patients.

VLE computer algorithm

Clinically inspired features versus generic image analysis features

Table 1 shows an overview of the performance of the four generic image analysis features and three clinically inspired features that were evaluated and compared by applying various widely used machine learning methods. The clinically inspired features achieved a better performance over the generic image analysis features for all evaluated machine learning methods. In figure 6 the optimal performances of the clinically inspired features are shown, which were derived after further optimization of the features during additional experiments.[18]

Performance computer algorithm versus VLE experts

Figure 6 shows the ROC curves of the computer algorithm for each of the clinically inspired features, compared with the performance of the VLE experts. The AUC was 0.95, 0.89 and 0.91 for the algorithm using the clinical features "layering and signal decay statistics", "signal intensity distribution" and "layering", respectively, compared to 0.81 for the VLE experts. For the VLE experts, a cut-off value of ≥ 8 points on the prediction score was associated with a sensitivity and specificity of 85% and 68% in the learning phase (20 images) and 83% and 71% in the validation phase (40 images), respectively. Derived from the ROC in figure 6, the sensitivity of the optimal feature "layering and signal decay statistics" was 90%, with a specificity of 93%. The features "layering" and "signal intensity distribution" showed a sensitivity and specificity of 83% and 93%, and 83% and 87%, respectively.

DISCUSSION

Our results show that a computer algorithm based on clinically derived features is capable to identify early BE neoplasia on ex vivo VLE images. In fact, compared to VLE experts that scored the same ex vivo VLE images, the algorithm performed even better in identifying neoplasia (AUC 0.89-0.95 vs 0.81 for experts). We envision that an objective and quantitative interpretation of VLE by a computer-aided algorithm, not hindered by inter-observer variability, will be an important tool to aid in the real-time VLE interpretation and detection of BE neoplasia during ongoing endoscopy.
Computer-aided diagnosis (CAD) is an emerging field in medical imaging in general and we foresee that computer aided detection will play an important role in the future developments of gastrointestinal endoscopy.[19] VLE scans contain a large amount of data. Computer-aided detection algorithms will be able to analyze all pixels in a VLE scan and can be trained to recognize abnormal patterns. In addition, algorithms could potentially reveal information in VLE scans that may not be

immediately apparent to the human eye. Thus, CAD may enable real-time analysis of VLE scans during Barrett's surveillance, red-flagging suspicious areas in the BE segment. With the new VLE system incorporating a laser marking tool these suspicious areas could subsequently be targeted by laser marks to allow VLE-targeted biopsies. In the future, this could cause a paradigm shift from the current random biopsy protocol towards VLE-targeted biopsies, saving both time and costs. Furthermore, algorithms could potentially aid in delineation of lesions with subtle margins or subsquamous extension. Recently, we demonstrated that a computer algorithm was able to accurately delineate early neoplastic lesions in BE using white light endoscopy images.[11] For detection of BE neoplasia on OCT images, Qi et al. developed a computer-aided detection algorithm to detect dysplasia in 106 OCT images from 13 BE patients.[12] In a successive study, a larger OCT dataset (690 OCT images from almost 100 biopsy sites) and improved image analysis methodology was used.[13] The image classification features were based on similar characteristics of neoplastic mucosa as in our study. In addition, use of multiple OCT frames in comparison to single image analysis was examined, resulting in a significantly improved sensitivity for classification of the images. Using ten frames per biopsy site, a sensitivity of 78%, specificity of 90% and accuracy of 84% was shown. The following differences can be observed when comparing this study with our study: First, although similar clinical features for neoplastic mucosa were used as a starting point, different image analysis methods were used for the final analysis. Second, a probe-based OCT system was used, with a small imaging surface, whereas VLE provides a circumferential imaging volume of 6-cm length. Third, although attention was paid to obtain optimal correlation of the OCT imaging site to the biopsy site, the quality of the OCT-histology correlation is not comparable to the high-quality one-to-one correlated VLE-histology database used in our study.[14] Fourth, images were categorized differently based on histological diagnosis: NDBE and LGD/HGD versus NDBE and HGD/EAC in our study.
The use of well-studied clinical features is an important basis for developing a robust computer algorithm for BE neoplasia detection. The clinically derived VLE features incorporated in our algorithm were identified through a multi-phased protocol with VLE expert scoring, using precise histological correlations.[6] The three identified VLE features for neoplasia were: 1) lack of layering, 2) VLE surface signal higher than subsurface signal, 3) presence of irregular glands. For the algorithm, only the first two VLE features were used. With the clinical input of these two features, three novel image analysis features were developed for the algorithm: 1) Layering, 2) Layering and signal decay statistics and 3) signal intensity distribution. For a range of commonly used machine learning methods, these clinically inspired features clearly outperformed all generic image analysis features that are widely used in several image analysis tasks (e.g. face recognition and traffic sign detection). Further investigation of the algorithm parameters showed that especially the top layers of the image were important for discrimination between neoplastic and NDBE tissue, confirming earlier clinical

findings.[6] This reinforces that the algorithm features connect well to the clinical input (layering). An explanation for the fact that the top layers perform better than the deeper layers in the algorithm, is that in the current VLE data the imaging quality decreases after a certain depth. This causes the signal-to-noise ratio to be lower in the deeper layers, which may potentially impede the identification of discriminative features, compared to the top layers. In ongoing studies, we are working from single image analysis towards multi-frame analysis. Simultaneous analysis of several adjacent slices is expected to further increase the performance of the algorithm. Eventually, full in vivo scan analysis in real-time during endoscopy will become possible. We expect that during the investigation of in vivo VLE scans, we may potentially identify additional features in the deeper esophageal wall layers, which could be implemented as image analysis features.

The VLE images used in this study were derived from a high-quality ex vivo VLE-histology database. However, ex vivo VLE data is limiting due to the presence of artifacts caused by endoscopic resection and histological processing. In addition, some structures (e.g., glands and blood vessels) may have collapsed, which could explain the limited amount of visible (irregular) glands in the ex vivo images. The results of a recent publication also suggested that in vivo VLE data may be easier to interpret. Eight VLE experts scoring 120 in vivo VLE images showed higher accuracy for the identification of neoplastic BE than in our study with ex vivo images.[6,20] However, scrutinizing the large amount of data that one VLE scan comprises will remain challenging for endoscopists. The number of images used in this feasibility study was small. Some image analysis methods could therefore not be assessed since these methods require a large dataset. The next step is to further develop and validate the algorithm on a large dataset of in vivo data, which is more representable for daily clinical practice. In vivo scans will be provided by future studies with the novel VLE laser marking system, which enables correlation of in vivo VLE images to histology.[21] With the availability of a larger database, future studies will also focus on deep learning (neural network); a successful method for various machine learning problems. The final aim will be automatic assessment of entire 3D VLE scans in real-time, enabling redflagging of neoplastic lesions on VLE during endoscopy. In the future, VLE could potentially enable replacement of numerous random biopsies by VLE-targeted biopsies during BE surveillance.

In conclusion, this is the first study in which a clinically inspired computer algorithm for BE neoplasia was developed, based on ex vivo VLE images with direct histological correlates. The algorithm showed good performance to detect BE neoplasia (AUC 0.95 with best performing features). This suggests that an automatic detection algorithm has the potential to assist endoscopists to detect early neoplasia on VLE. However, future studies on in vivo VLE scans are needed to further expand and validate the algorithm on in vivo data.

ACKNOWLEDGEMENTS

This study was supported by an unrestricted grant from NinePoint Medical (NinePoint Medical Inc., Bedford, MA, USA).

REFERENCES

1. Pohl H, Sirovich B, Welch HG. Esophageal adenocarcinoma incidence: are we reaching the peak? Cancer Epidemiol Biomarkers Prev. 2010 Jun;19(6):1468–70.

2. Fitzgerald RC, di Pietro M, Ragunath K, Ang Y, Kang J-Y, Watson P, et al. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. Gut. 2014 Jan;63(1):7–42.

3. Shaheen N, Falk G, Iyer P, Gerson L. ACG Clinical Guideline: Diagnosis and Management of Barrett's esophagus. Am J Gastroenterol. 2015;108(8):1238–49.

4. Suter MJ, Vakoc BJ, Yachimski P, Shishkov M, Lauwers GY, Mino-Kenudson M, et al. Comprehensive microscopy of the esophagus in human patients with Optical Frequency Domain Imaging. Array, editor. Gastrointest Endosc. 2008;68(4):745–53.

5. Wolfsen HC, Sharma P, Wallace MB, Leggett C, Tearney G, Wang KK. Safety and feasibility of volumetric laser endomicroscopy in patients with Barrett's esophagus (with videos). Gastrointest Endosc. 2015;3:1–10.

6. Swager A, Tearney G, Leggett C, van Oijen M, Meijer S, Weusten B, et al. Identification of Volumetric Laser Endomicroscopy features predictive for early neoplasia in Barrett's esophagus using high-quality histological correlation. Gastrointest Endosc. 2016;Epub ahead of print.

7. Leggett CL, Gorospe EC, Chan DK, Muppa P, Owens V, Smyrk TC, 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 May;83(5):880–888.e2.

8. Swager A, van Oijen M, Tearney G, Leggett C, Meijer S, Bergman J, et al. How Good are Experts in Identifying Early Barrett's Neoplasia in Endoscopic Resection Specimens Using Volumetric Laser Endomicroscopy? Gastroenterology. 2016;150(4):S628.

9. Swager A, van Oijen M, Tearney G, Leggett C, Meijer S, Bergman J, et al. How Good Are Experts in Identifying Endoscopically Visible Early Barrett's Neoplasia on in vivo Volumetric Laser Endomicroscopy? Gastrointest Endosc. 2016;83(5):AB573.

10. Trindade AJ, George BJ, Berkowitz J, Sejpal D V, McKinley MJ. Volumetric laser endomicroscopy can target neoplasia not detected by conventional endoscopic measures in long segment Barrett's esophagus. Endosc Int open. 2016;4(3):E318-22.

11. Sommen F Van Der, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BLAM, et al. Computer-aided detection of early neoplastic lesions in Barrett ' s esophagus. Endoscopy. 2016;48(7):617–24.

12. Qi X, Sivak M V, Isenberg G, Willis JE, Rollins AM. Computer-aided diagnosis of dysplasia in Barrett's esophagus using endoscopic optical coherence tomography. J Biomed Opt. 2006;11(4):44010.

13. Qi X, Pan Y, Sivak M V, Willis JE, Isenberg G, Rollins AM. Image analysis for classification of dysplasia in Barrett's esophagus using endoscopic optical coherence tomography. Biomed Opt Express. 2010;1(3):825–47.

14. Swager A, Boerwinkel DF, de Bruin DM, Weusten BL, Faber DJ, Meijer SL, et al. Volumetric

laser endomicroscopy in Barrett's esophagus: A feasibility study on histological correlation. Dis Esophagus. 2016;29:505–12.

15. Yun SH, Tearney GJ, Vakoc BJ, Shishkov M, Oh WY, Desjardins AE, et al. Comprehensive volumetric optical microscopy in vivo. Nat Med. 2006;12:1429–33.

16. Yun S, Tearney G, De Boer J, Iftimia N, Bouma B. High-speed optical frequency-domain imaging. Opt Express. 2003;11(22):2953–63.

17. Vakoc BJ, Shishko M, Yun SH, Oh W, Tearney GJ, Bouma BE. Comprehensive esophageal microscopy by using optical frequency-domain imaging (with video). Gastrointest Endosc. 2007;65(6):898–905.

18. Klomp SR, van der Sommen F, Swager A, Zinger S, Schoon EJ, Curvers WL, et al. Evaluation of Image Features and Classification Methods for Barrett's Cancer Detection Using VLE Imaging. To be Present SPIE Med Imaging, Febr 2017, Orlando, FL, USA.

19. Curvers WL, Bergman JJ. A new paradigm shift in endoscopy: From interpretation to automated image analysis? Gastrointest Endosc. 2016;83(1):115–6.

20. Trindade AJ, Inamdar S, Smith MS, Chang KJ, Leggett CL, Lightdale CJ, et al. Volumetric laser endomicroscopy in Barrett's esophagus: interobserver agreement for interpretation of Barrett's esophagus and associated neoplasia among high-frequency users. Gastrointest Endosc. 2016 Nov 26;Epub ahead of print.

21. Suter MJ, Gora MJ, Lauwers GY, Arnason T, Sauk J, Gallagher K a., et al. Esophageal-guided biopsy with volumetric laser endomicroscopy and laser cautery marking: a pilot clinical study. Gastrointest Endosc. 2014 Jan;2:1–11.

22. Rodriguez-Diaz E, Singh SK. Computer-Assisted Image Interpretation of Volumetric Laser Endomicroscopy in Barrett's Esophagus. Gastroenterology. 2015;148(4):S91–S92.

FIGURES AND TABLES

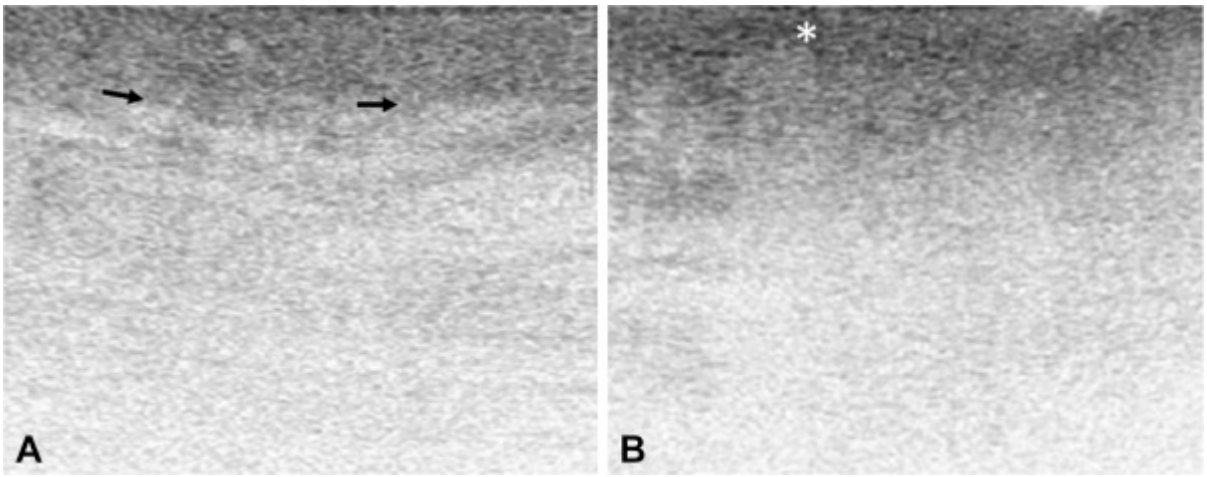

Fig. 1: Examples of normalized VLE images, containing NDBE (A) and neoplasia (EAC) (B) on corresponding histology. In (A) a layered pattern is visible (arrows). In (B) lack of layering is visible and a higher surface than subsurface signal (asterisk).

Fig. 2: Commercial Nvision VLE Imaging System.

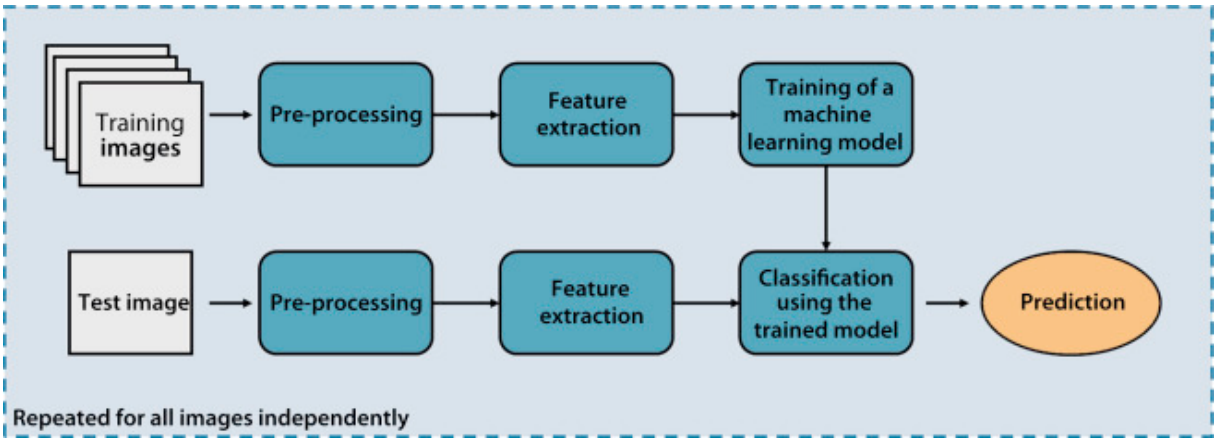

Fig. 3: General framework for the VLE computer algorithm, where the system is trained with a set of images and evaluated on a new unseen image. The framework consists of the following steps: 1) Preprocessing: the region of interest containing only tissue was selected (non-tissue features like air or cork where the specimens were pinned on, were excluded). Thereafter, each image was standardized to a height of 400 pixels. 2) Feature extraction: Information in the image can be quantified into numbers in order to derive informative values (features) from the data. 3) Machine learning: this is a general term for learning a computer algorithm to recognize patterns in data.

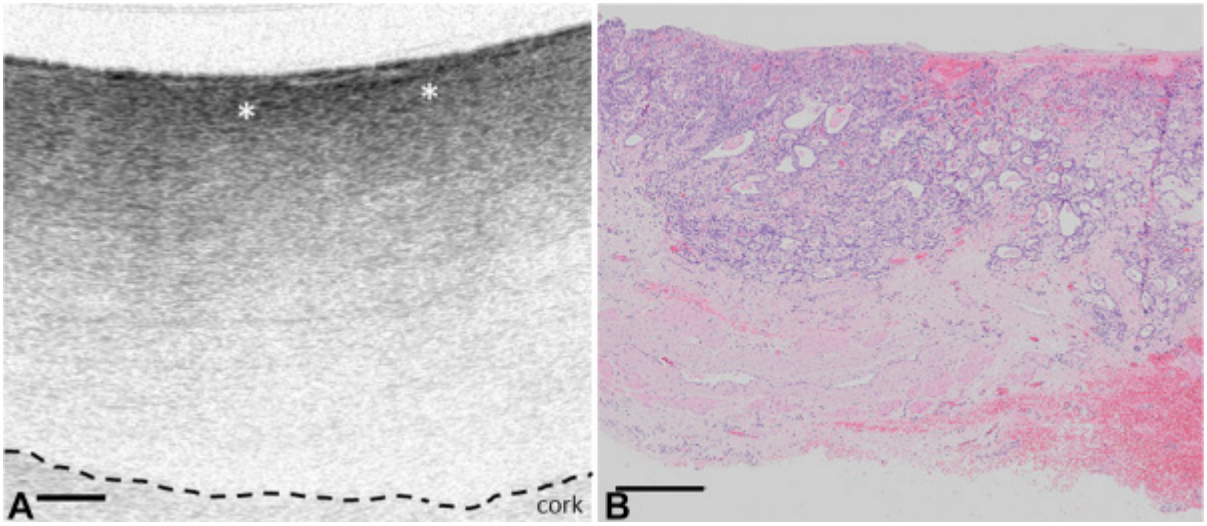

Fig. 4: Example of a VLE image showing lack of layering and a surface VLE signal that is higher than subsurface signal (asterisks) (A), corresponding with EAC on histology (B). The dotted line in (A) represents separation between cork and tissue. *EAC*, early adenocarcinoma; *VLE*, volumetric laser endomicroscopy. Scale bars represent 500 µm.

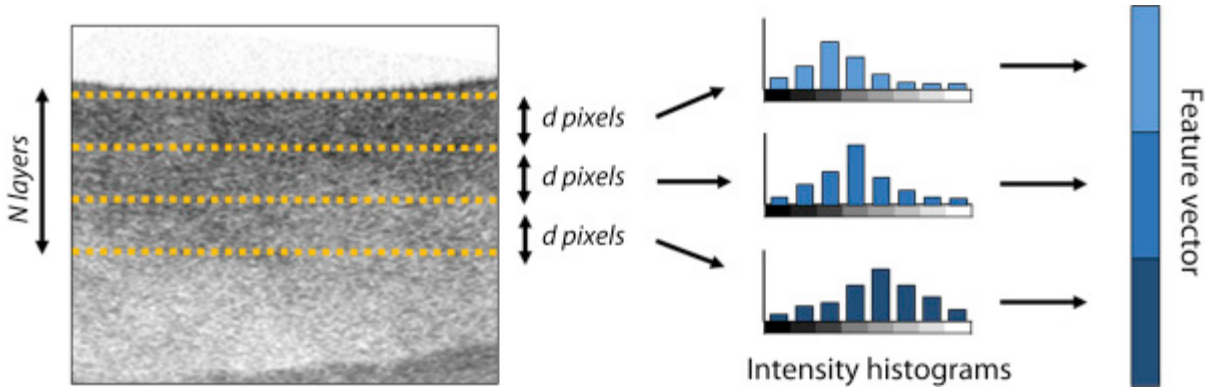

Fig. 5: The first novel clinically inspired algorithm feature termed “layering” was constructed by calculating a VLE signal intensity histogram using 8 intensity categories over a number of layers. Based on the intensity histograms, a feature vector (consisting of an array of numbers) was constructed, which was used as quantitative input for the machine learning methods. This figure shows a schematic of the construction of the feature vector. For constructing the feature vectors, histograms were computed from the first N layers of d pixels and combined in one array of numbers.

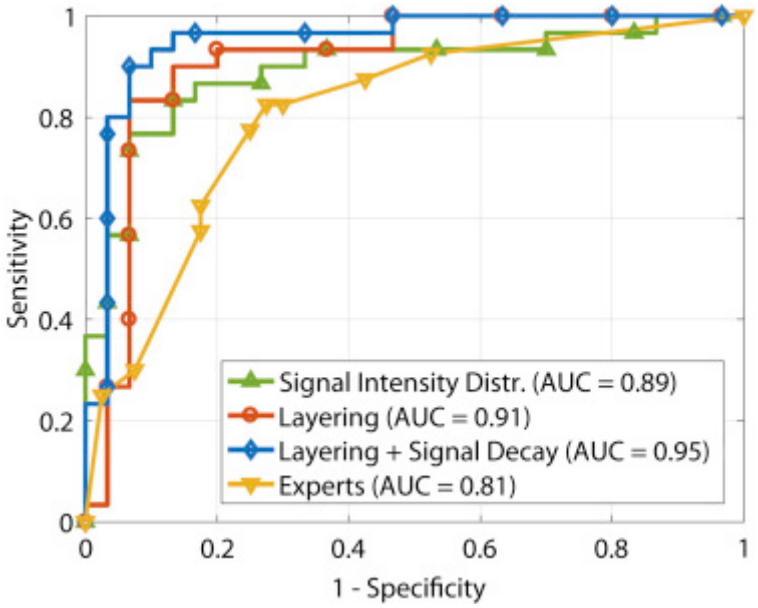

Fig. 6: ROC curve for the three proposed clinically inspired features compared to the ROC for the experts.

Machine learning methods	Generic image analysis features				Clinically inspired features		
	1 (Gray-Level Co-occurrence Matrices)	**2** (Local Binary Patterns)	**3** (Histogram of Oriented Gradients)	**4** (Wavelet transform[22])	Layering	Signal intensity distribution	Layering and signal decay statistics
1 (Support Vector Machine)	0.65	0.70	0.72	0.74	0.80	**0.82**	0.81
2 (Discriminant Analysis)	0.59	0.70	0.70	0.69	**0.75**	0.74	0.74
3 (AdaBoost)	0.68	0.64	0.61	0.70	0.84	0.90	**0.91**
4 (Random Forest)	0.66	0.64	0.72	0.63	0.75	**0.83**	0.81
5 (K-Nearest Neighbors)	0.63	0.68	0.64	0.64	**0.80**	0.77	0.61
6 (Naïve Bayes)	0.53	X	X	0.67	**0.73**	X	**0.73**
7 (Linear Regression)	0.59	0.57	X	0.70	**0.72**	0.62	**0.72**
8 (Logistic Regression)	0.57	0.61	X	0.65	0.71	X	**0.73**

Table 1: Comparison of generic image features with the clinically inspired image features for several common machine learning methods. The numbers indicate the area under the receiver operator characteristics curve (AUC). Superior performance is highlighted in gray.