

Evaluation of quantitative image analysis criteria for the high-resolution microendoscopic detection of neoplasia in Barrett's esophagus

Timothy J. Muldoon

Nadhi Thekkek

Darren Roblyer

Rice University
Department of Bioengineering
6100 Main Street
Houston, Texas 77005

Dipen Maru

The University of Texas
M.D. Anderson Cancer Center
Department of Pathology
1515 Holcombe Boulevard
Houston, Texas 77030

Noam Harpaz

Mount Sinai Medical Center
Division of Pathology
1 Gustave L. Levy Place
New York, New York 10029

Jonathan Potack

Sharmila Anandasabapathy

Mount Sinai Medical Center
Division of Gastroenterology
1 Gustave L. Levy Place
New York, New York 10029

Rebecca Richards-Kortum

Rice University
Department of Bioengineering
6100 Main Street
Houston, Texas 77005

1 Introduction

The incidence of esophageal adenocarcinoma (EAC) has increased sixfold in the United States over the past 30 years.¹ Adenocarcinoma of the distal esophagus has one of the poorest overall five-year survival rates of all cancers, with reported five-year survival rates of 14% and 0% for stage III and stage IV disease, respectively.² Surgical treatment for locally advanced EAC carries significant morbidity and mortality. Early detection of disease is vital to improve long-term survival rates, facilitate endoscopic therapy, and improve the overall quality of life of affected patients.

The current method of surveillance of Barrett's esophagus involves endoscopic white-light examination with four-

Abstract. Early detection of neoplasia in patients with Barrett's esophagus is essential to improve outcomes. The aim of this *ex vivo* study was to evaluate the ability of high-resolution microendoscopic imaging and quantitative image analysis to identify neoplastic lesions in patients with Barrett's esophagus. Nine patients with pathologically confirmed Barrett's esophagus underwent endoscopic examination with biopsies or endoscopic mucosal resection. Resected fresh tissue was imaged with fiber bundle microendoscopy; images were analyzed by visual interpretation or by quantitative image analysis to predict whether the imaged sites were non-neoplastic or neoplastic. The best performing pair of quantitative features were chosen based on their ability to correctly classify the data into the two groups. Predictions were compared to the gold standard of histopathology. Subjective analysis of the images by expert clinicians achieved average sensitivity and specificity of 87% and 61%, respectively. The best performing quantitative classification algorithm relied on two image textural features and achieved a sensitivity and specificity of 87% and 85%, respectively. This *ex vivo* pilot trial demonstrates that quantitative analysis of images obtained with a simple microendoscope system can distinguish neoplasia in Barrett's esophagus with good sensitivity and specificity when compared to histopathology and to subjective image interpretation. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3406386]

Keywords: endoscopy; biomedical optics; microscopy.

Paper 09326SSR received Aug. 1, 2009; revised manuscript received Jan. 11, 2010; accepted for publication Feb. 8, 2010; published online Apr. 28, 2010.

quadrant biopsy, a procedure that has been shown to miss neoplasia in up to 57% of cases.³⁻⁶ Because of these limitations, high-resolution technologies that can serve as an adjunct to white-light endoscopy have been proposed to increase diagnostic accuracy for the detection of high-grade dysplasia (HGD) or early cancer. These technologies include confocal microendoscopy and endocytoscopy. Of these, confocal endomicroscopy has shown the highest accuracy to date based on single-center evaluation with an experienced microendoscopist interpreting the *in vivo* images.⁷ Unfortunately, such real-time interpretation requires considerable training and is subject to intraobserver variability.⁸⁻¹⁰ For broader clinical application, particularly outside a university setting, the application of image-analysis software using well-defined image classification algorithms has the potential to reduce subjectivity and improve diagnostic accuracy.

Address all correspondence to: Rebecca Richards-Kortum, Rice University, Department of Bioengineering, 6100 Main Street, Houston, Texas 77005. Tel: 713-348-5869; Fax: 713-348-5877; E-mail: rkortum@rice.edu

Quantitative image analysis, based on computational analysis of textural features within a digital image, offers an objective means to improve the consistency and speed of diagnosis at the point of care. Image texture features, based on the spatial distribution and organization of pixels, can be computed rapidly. Textural features, such as entropy, frequency content, and pixel pair correlation values, may not be apparent to a human observer, but they relate important information about the structure within an image. Images from a dysplastic region, for example, would be more likely to have a lower pixel pair correlation value and a higher entropy value because of the decreasing orderly arrangement of cells (and therefore corresponding image pixels) when compared to normal glandular tissues. Since there are a large number of easily calculated textural features, searching many features at once can identify a single feature or pair of features that enable high-fidelity classification of the image data. By limiting the number of features calculated at the time of image acquisition, rapid diagnostics are possible at the point of care. With the increasing availability and power of inexpensive and portable computers, on-site quantitative analysis and classification techniques could be applied to a wide array of high-resolution imaging schemes.

To extend the benefits of high-resolution imaging to a broader patient population, and to create a more objective means of evaluating optical biopsy data, we developed a quantitative image analysis and classification algorithm to analyze data from a high-resolution microendoscope (HRME) device. The HRME is capable of producing images of the esophageal mucosa at subcellular resolution without the need for expensive optics or scanning electronics. Indeed, the total cost of the HRME components is approximately \$2500. In a previous proof-of-principle study, we demonstrated that images acquired with the HRME can distinguish neoplastic [high-grade dysplasia (HGD) or cancer] from non-neoplastic [intestinal metaplasia or low-grade dysplasia (LGD)], Barrett's mucosa based on qualitative image assessment.¹¹

The goal of this study was to develop quantitative image analysis criteria for HRME images to discriminate neoplastic from non-neoplastic esophageal mucosa. We compared the accuracy of discrimination achieved using visual interpretation by experienced pathologists or gastroenterologists trained in the analysis of the optical images to that achieved with computer-aided classification algorithms based on quantitative analysis of image textural features. Histopathology was considered the gold standard.

2 Methods

2.1 Patients

Patients over 18 years of age with a previous diagnosis of Barrett's esophagus were asked to participate in the study. Informed consent was obtained from all study participants, and the study was reviewed and approved by the Institutional Review Boards at the University of Texas M. D. Anderson Cancer Center, Rice University, and The Mount Sinai Medical Center. Subjects underwent conventional endoscopy with standard four-quadrant biopsy surveillance; a subset of patients (those with endoscopically suspected neoplasia) underwent endoscopic mucosal resection (EMR). Following resec-

tion, biopsies or EMR specimens were imaged with the HRME device.

A solution of proflavine (Sigma-Aldrich) dissolved in water at a concentration of 0.01% (w/v) was prepared prior to performing imaging. The contrast agent solution was directly applied to the epithelial surface of the resected tissue with a dropper, and imaging with the HRME device was performed immediately. The application of proflavine does not discolor the tissue surface, and is not detectable in tissue slides prepared using standard histopathology processing and hematoxylin and eosin (H&E) staining (Fig. 1).

After imaging, the tissue was returned for standard histopathology processing, and slides were later reviewed by a single, expert gastrointestinal pathologist (D.M.) blinded to the results of the HRME imaging. Each measurement site used in this study was correlated to a histopathology-confirmed diagnostic category: Barrett's intestinal metaplasia (IM), low-grade dysplasia (LGD), high-grade dysplasia (HGD), or esophageal adenocarcinoma (EAC).

2.2 Imaging System

The high-resolution microendoscope (HRME) device has been previously described in detail.¹² Briefly, images are acquired with this device by placing the tip of the fiber bundle image guide into direct contact with the epithelial surface of the tissue. Excitation light from a blue LED with a center wavelength of 455 nm is delivered through the fiber bundle. The fluorescence emission from the topically applied fluorescent contrast agent, proflavine, is collected through the fiber bundle and focused onto a CCD camera, and a digital image is stored for future processing and analysis. The HRME system has a circular field of view with a diameter of 750 microns; the lateral spatial resolution of the system is approximately 4 microns, and images are displayed at 4 frames per second.

2.3 HRME Image Analysis

Digital HRME images were reviewed to determine whether the endoscope tip was in contact with the tissue surface or whether the probe tip moved during image acquisition. Images showing such artifacts were discarded and not used in subsequent analyses.

Digital HRME images were reviewed by two expert pathologists and two expert gastroenterologists, already familiar with microendoscopy. Prior to reviewing the entire set of images, reviewers were shown a subset of 16 images, labeled with the corresponding histopathologic diagnosis of IM, LGD, HGD, and EAC. This training set included 8 images collected from sites with a diagnosis of IM or LGD (non-neoplastic), and 8 images from sites diagnosed with neoplasia (HGD or EAC). Reviewers were then shown the complete set of images in a randomized order and asked to score each image as either "neoplastic" or "non-neoplastic," where "non-neoplastic" corresponds to a pathologic diagnosis of IM or LGD, and "neoplastic" corresponds to HGD or EAC. Results of visual image interpretation were compared to the histopathology-confirmed diagnosis at each site; sensitivity and specificity were calculated for each observer.

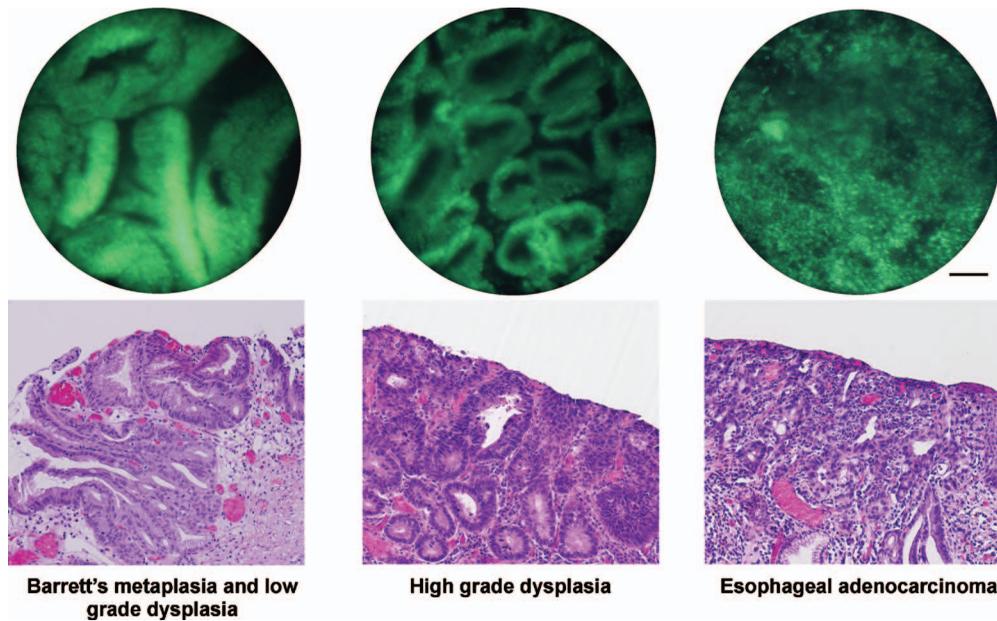


Fig. 1 Illustration of HRME images (top row) compared to H&E images (bottom row). From left to right, diagnostic categories are: Barrett's metaplasia and low-grade dysplasia (LGD), high-grade dysplasia (HGD), and esophageal adenocarcinoma (EAC). All images are sized to the same scale; scale bar represents 100 μm .

2.4 Image Classification Algorithm

In addition to subjective image interpretation, we explored the diagnostic ability of quantitative image analysis. For each image, 59 distinct features were computed (Table 1). First-order statistical features (variance, entropy, etc.) were calculated directly from the raw pixel values. A gray-level co-occurrence matrix (GLCM) with pixel offsets ranging from 1 to 10 was used to calculate additional textural feature groups (correlation, contrast, homogeneity, and energy).¹³ Each GLCM feature group contained 10 distinct features, corresponding to each pixel offset. To detect nuclear features, an extended regional maximum transform was applied to the image. Voronoi tessellations were calculated from the centroids of the nuclear features to calculate internuclear distances.¹⁴ A Fourier transform was applied to each image to calculate the power spectrum; this was divided into 10 partitions to represent the frequency components of the image. The contribution of each partition is represented as a fraction of the total power spectrum.^{15,16}

A diagnostic algorithm was developed to classify each image as non-neoplastic or neoplastic using these image features as input. Histopathology again was used as the gold standard; sites with a pathologic diagnosis of Barrett's metaplasia or Barrett's metaplasia with low-grade dysplasia were considered to be non-neoplastic, while sites with a pathologic diagnosis of Barrett's metaplasia with high-grade dysplasia or esophageal adenocarcinoma were considered to be neoplastic. The classifier was based on two-class, linear discriminant analysis; a sequential forward feature-selection algorithm initially was used to select the best performing subset of up to 10 image features to classify the data. Initially, the best performing single feature was identified, and then subsequent features were selected that gave maximum performance when combined with previously selected features. The algorithm, was

developed using fivefold cross-validation; each measurement site was initially randomly assigned to one of five groups. Four-fifths of the data were then used to train the linear classification algorithm, and the remaining one-fifth of the data were used to test the algorithm. This cycle was repeated four-additional times so that the algorithm was tested using data from each site. Performance was monitored by calculating the area under the curve (AUC) of the classifier.

Alternatively, a three-class diagnostic algorithm was developed to classify each image as: (1) Barrett's intestinal metaplasia or Barrett's low-grade dysplasia, (2) Barrett's high-grade dysplasia, and (3) adenocarcinoma. A categorical tree-based classifier was used. This algorithm was allowed to choose inputs from the entire feature set.^{17,18} The tree-based classifier was pruned to three terminal nodes to avoid over-training; two features were selected to perform this step. The predicted classification results from each classifier were then compared to the actual histopathology for each site.

3 Results

3.1 Patients and Measurement Sites

Nine subjects were enrolled in the study. Endoscopic mucosal resection (EMR) specimens were obtained from six of these patients, and biopsy specimens were obtained from the remaining three patients. Images were obtained from 139 unique sites; images from 128 of these sites passed the quality control (QC) review and were used for further analysis (Table 2). Figure 1 shows representative HRME images (top row) and corresponding histopathology (bottom row) of metaplasia/LGD, HGD, and EAC, respectively. All images are at the same scale for comparison. Large, well-organized glands can be seen in the metaplasia/LGD Barrett's case, while smaller glands with disrupted borders are visible in the

Table 1 Quantitative image features. Image features are listed in decreasing order of diagnostic performance when used as a single input feature for linear discriminant analysis.

Feature	Explanation
Correlation (10 features)	Pixel neighbor correlation over the entire image
Standard deviation	Standard deviation of grayscale values
Variance	Variance of pixel grayscale values
Energy (10 features)	Sum of squares in gray-level co-occurrence matrix (GLCM)
Frequency (10 features)	Frequency distribution of pixel values
Entropy	Statistical measurement of randomness of grayscale values
Mean nuclear separation distance	Mean nuclear separation as calculated by Voronoi tessellation
Std. dev. nuclear separation distance	Standard deviation of nuclear separation as calculated by Voronoi tessellation
Nuclei per unit area	Number of nuclei detected divided by area of region of interest
Kurtosis	Measure of the flatness of the pixel value distributions
Skewness	Measure of the symmetry of the pixel value distribution
Contrast (10 features)	Measure of pixel intensity compared to its neighbors over the entire image
Homogeneity (10 features)	Closeness of the distribution of the GLCM elements to the diagonal
Mean minimum nuclear separation distance	Average minimum nuclear separation as calculated by Voronoi tessellation

Table 2 Number of sites imaged by histologic diagnosis.

Diagnostic category	Histologic diagnosis	Number of sites imaged
Non-neoplastic	IM	6
	LGD	53
Neoplastic	HGD	26
	EAC	43
	Total	128

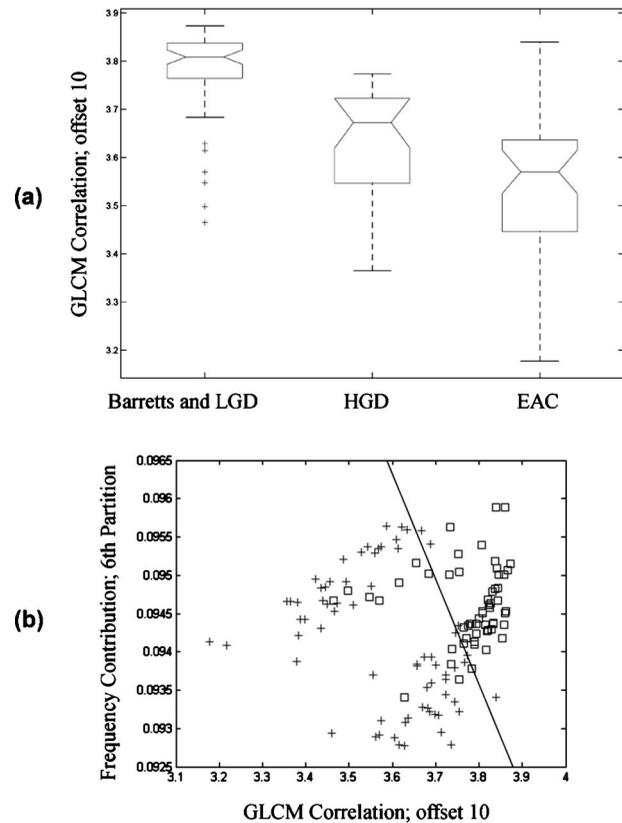


Fig. 2 (a) Box and whiskers plot showing the GLCM correlation by diagnostic category. The central line in each box represents the mean, while the top and bottom edges represent the 25th and 75th percentile. The notches in the boxes represent a 95% confidence interval for the mean. (b) Features values for top two performing features by measurement site. Non-neoplastic sites are plotted as squares; neoplastic sites are plotted as crosses. The decision line is shown.

HGD case. The image of EAC shows extreme disruption of glandular organization and crowded, abnormal cells.

3.2 Subjective Image Interpretation

Subjective scoring of the images by two expert gastroenterologists achieved an average sensitivity of 87% and an average specificity of 53% for neoplasia, with a kappa statistic of 0.25, indicating fair agreement. The two expert pathologists achieved an average sensitivity of 87% and an average specificity of 68% for neoplasia, with a kappa statistic of 0.27, indicating fair agreement. The combined average performance of all four reviewers was 87% sensitivity and 61% specificity.

3.3 Quantitative Classification Algorithm

Quantitative image features were ranked according to their diagnostic ability. Image features in Table 1 are listed in order of descending area under the receiver operating characteristic (ROC) curve for the single-feature, two-class linear discriminant analysis classifier. The single best performing feature was found to be the GLCM correlation with a pixel offset of 10. Figure 2(a) depicts a box plot illustrating the average GLCM correlation values for samples diagnosed histologically as non-neoplastic samples (IM or LGD), HGD, and

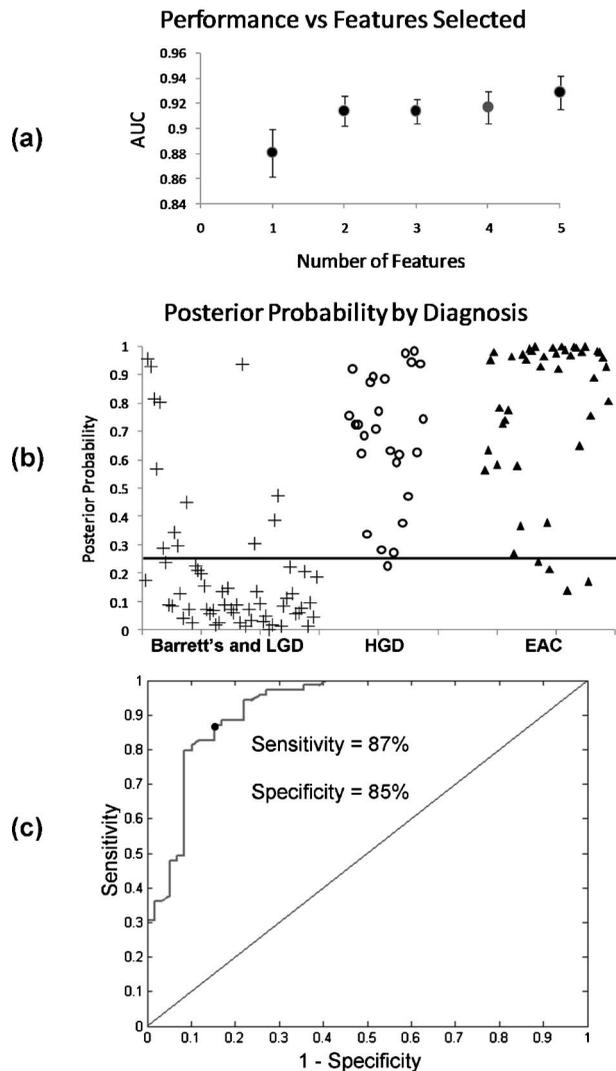


Fig. 3 Classification performance. (a) Calculated classification performance versus number of features used; note the plateau at two features. (b) Scatter plot of calculated posterior probability generated by a two-feature linear discriminant analysis (LDA) algorithm. The decision line is shown at 0.24. (c) ROC curve of calculated classification performance using the two-class, two-feature LDA. The Q-point is shown, corresponding to a sensitivity of 87% and a specificity of 85%.

EAC. On average, the GLCM correlation level is lower in sites with HGD or EAC when compared to non-neoplastic sites.

The top performing combination of two features was found to be the GLCM correlation value at an offset of 10, and the frequency contribution at an offset of 6; this is the relative contribution to the total power spectrum coming from the frequencies over the sixth partition from the Fourier transform. Figure 2(b) shows a scatter plot of these two features for each of the 128 sites in the data set. Non-neoplastic samples, with a pathologic diagnosis of Barrett’s with or without LGD, are shown as squares; neoplastic samples, with a pathologic diagnosis of HGD or EAC, are shown as crosses. The decision line associated with the two-class linear discriminant classifier is shown as a straight line on the plot.

Figure 3(a) shows the area under the curve as a function of

Table 3 Summary of the performance of the tree-based three-class classifier.

	IM/LGD by histology	HGD by histology	EAC by histology
IM/LGD predicted	79.3%	7.4%	2.3%
HGD predicted	8.6%	70.4%	4.7%
EAC predicted	12.1%	22.2%	93.0%

the number of image features selected in the linear discriminant algorithm. The AUC increases from one to two features and then reaches a plateau. Figure 3(b) shows a scatter plot of the posterior probability that each site is neoplastic (HGD or EAC) as calculated by the linear classifier using two input features; samples are grouped by histopathologic diagnosis. Figure 3(c) shows the ROC curve for the linear discriminant classifier based on these two features. The AUC is 0.92, and the sensitivity and specificity at the Q-point are 87% and 85%, respectively.

A tree-based classifier was developed to classify samples into three groups (non-neoplastic, HGD, and cancer) using automated feature selection. The features chosen by this algorithm were again GLCM correlation and frequency contribution. Table 3 summarizes the performance of the three-class algorithm; vertical columns add to 100%, to indicate the proportion of predicted measurements that were placed into the correct category. The tree-based classifier performed very well in distinguishing the non-neoplastic cases (79.3%) and when predicting HGD and EAC (70.4% and 90.3%, respectively).

4 Discussion

In this study, we demonstrated the use of a simple, low-cost, portable, high-resolution microendoscopy system to distinguish between clinically significant grade of Barrett’s esophagus using subjective visual interpretation and objective quantitative image analysis and classification. Subjective analysis of the images by expert gastroenterologists achieved average sensitivity and specificity of 87% and 53%, respectively, while expert pathologists achieved average sensitivity and specificity of 87% and 68%, respectively. Image classification algorithms were created by analysis of key textural features within the images, the most important features being frequency content, as calculated by a discrete fast Fourier transform, and pixel pair correlation, as calculated by a gray-level co-occurrence matrix. The objective classification algorithm was able to distinguish between neoplastic and non-neoplastic cases with a sensitivity of 87% and a specificity of 85%.

The results of this pilot study suggest that this technique may be useful to regions without highly trained expert personnel or extensive biomedical infrastructure. The low cost of the device (roughly US \$2,500) and the wide availability of low-cost computers may facilitate the distribution of the HRME and the quantitative image analysis and classification algorithm to regions beyond tertiary care centers and university hospitals.¹¹

Subjective scoring of the HRME images identified that significant intra-observer variability exists. While the overall accuracy is comparable to the results achieved by the computational diagnostic algorithm presented here, the subjective interpretation result demonstrates that even among highly trained endoscopists (familiar with microendoscopy) and expert gastrointestinal (GI) pathologists, disagreement can occur. Unfamiliarity with the image modality presented here by the HRME may be a confounding factor, as can bias from extensive experience with more conventional modalities, such as histopathology. The objective algorithm presented here may be able to aid clinicians in making diagnostic decisions at the point of care, while reducing intra-observer variability.

As with any small pilot study, a larger sample size will be required to optimize and evaluate the performance of the classification algorithm presented here. Results with the relatively small number of patients (9) and measurement sites (128), as well as the distribution of diagnoses measured (Table 2) may not reflect the HRME probe's performance in a screening setting. The non-neoplastic sites examined here included both intestinal metaplasia and low-grade dysplasia; however, the number of sites with low-grade dysplasia (53) was much greater than the number with intestinal metaplasia (6). While this is expected in the high-risk population participating in this study, it may induce bias in the accuracy rates reported in this feasibility study. In a normal screening population, the proportion of non-neoplastic sites with intestinal metaplasia only is expected to be higher than that encountered here.¹⁹ We are evaluating the performance of the algorithms developed here in a larger number of subjects in a normal screening population of average risk. A separate, independent validation data set is the most robust method to verify the performance of the HRME device and classification algorithm. However, in the absence of an independent validation set, cross-validation is recognized as a helpful tool to guard against overtraining when estimating the performance of a classifier.²⁰

An additional limitation of this trial is the HRME probe's inability to image subsurface regions. While the incidence of high-risk submucosal or occult EAC is still subject to controversy, inspection of subsurface regions remains an important consideration to clinicians.²¹ The current HRME device yields images by placing the tip of the device into direct contact with tissue, limiting its image acquisition to the superficial epithelium. However, the fiber-optic probe is small enough to pass through the lumen of a 16-gauge needle; the needle could be used to physically penetrate into deeper layers of the epithelium, allowing the fiber bundle to image these areas. This technique has been successfully demonstrated in a mouse model.¹²

The quantitative algorithm described in this paper could be applied to any high-resolution imaging system capable of digital image acquisition. In particular, confocal imaging, which has shown much promise as an *in situ* high-resolution imaging device, could benefit from objective image analysis. While the same features that were demonstrated to be diagnostic for this study (frequency content and pixel pair correlation) may not be applicable to all high-resolution imaging systems, a similar textural feature search could be employed. The new algorithm may then be used to augment physician diagnosis during a procedure.

High-resolution imaging of the gastrointestinal epithelium offers clinicians a means to inspect the histologic features of suspicious lesions and post-resection margins in real time during an endoscopic procedure. Such information can be used to guide the selection of sites for biopsy to improve the diagnostic yield during screening endoscopy.²² Traditional confocal laser endoscopy (CLE) typically requires the endoscopist to interpret images during the procedure. This subjective interpretation is likely highly dependent on clinician training and is variable by its subjective nature. A more objective means of image interpretation may help guide the endoscopist's decision as to the classification of a lesion. Such objective image analysis has been applied in other high-resolution imaging studies; Becker et al. demonstrated that computer-aided diagnosis of mucosal pit patterns but a fiber bundle confocal system has been shown to be effective in distinguishing neoplastic from non-neoplastic Barrett's tissue.²³

While the small field of view (750 microns) of the HRME device makes surveillance of the entire esophageal mucosa impractical, the HRME could be used as a tool to investigate suspicious sites during an endoscopic procedure. Standard white-light endoscopy has a field of view of several centimeters and can be used to identify regions of interest that require high-resolution interrogation with the HRME probe. In addition, other wide-field imaging techniques are becoming available, such as narrow-band imaging (NBI) and autofluorescence (AF) imaging, which could be useful in uncovering occult premalignant conditions.²⁴ Such techniques are susceptible to confounding factors, such as inflammation, which can be misclassified as dysplasia. High-resolution imaging of these regions may be able to correctly differentiate benign inflammatory processes from precancerous lesions, resulting in improved biopsy localization and higher diagnostic yield.

The image processing techniques presented in this study can uncover features in images that are clinically important, but that may not be apparent to the clinician performing the procedure. In addition, quantitative image analysis and classification techniques require relatively minimal computing time, which is useful in a busy clinical setting. This feasibility study demonstrates that images acquired with a low-cost, high-resolution microendoscope can provide objective discrimination of neoplastic and non-neoplastic sites in the esophagus with good accuracy relative to histology. Prospective evaluation of this approach is warranted to determine the clinical performance in a low-risk screening setting.

Acknowledgments

The authors would like to acknowledge funding from NIH Grant Nos. 1R01EB007594-01 and 1R01CA103830-05. One of the co-authors (R.R.K.) has a small ownership interest in Remicalm, Inc., which has licensed related technology from the University of Texas at Austin and Rice University.

References

1. H. Pohl and H. G. Welch, "The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence," *J. Natl. Cancer Inst.* **97**(2), 142–146 (2005).
2. G. Portale, J. A. Hagen, J. H. Peters, L. S. Chan, S. R. DeMeester, T. A. Gandamihardja, and T. R. DeMeester, "Modern 5-year survival of resectable esophageal adenocarcinoma: single institution experience with 263 patients," *J. Am. Coll. Surg.* **202**(4), 588–596; discussion 596 (2006).

3. M. Vieth, C. Ell, L. Gossner, A. May, and M. Stolte, "Histological analysis of endoscopic resection specimens from 326 patients with Barrett's esophagus and early neoplasia," *Endoscopy* **36**(9), 776–781 (2004).
4. W. L. Curvers and J. J. Bergman, "Multimodality imaging in Barrett's esophagus: looking longer, seeing better, and recognizing more," *Gastroenterology* **135**(1), 297–299 (2008).
5. E. S. Dellon and N. J. Shaheen, "Does screening for Barrett's esophagus and adenocarcinoma of the esophagus prolong survival?," *J. Clin. Oncol.* **23**(20), 4478–4482 (2005).
6. J. W. van Sandick, J. J. B. van Lanschot, B. W. Kuiken, G. N. J. Tytgat, G. J. A. Offerhaus, and H. Obertop, "Impact of endoscopic biopsy surveillance of Barrett's oesophagus on pathological stage and clinical outcome of Barrett's carcinoma," *Gut* **43**(2), 216–222 (1998).
7. R. Kiesslich, L. Gossner, M. Goetz, A. Dahlmann, M. Vieth, M. Stolte, A. Hoffman, M. Jung, B. Nafe, P. R. Galle, and M. F. Neurath, "In vivo histology of Barrett's esophagus and associated neoplasia by confocal laser endomicroscopy," *Clin. Gastroenterol. Hepatol.* **4**(8), 979–987 (2006).
8. M. Vieth and R. Kiesslich, "Distinction of high-grade intraepithelial neoplasia and invasive Barrett's adenocarcinoma," *Esophagus* **3**(4), 171–176 (2006).
9. Y. T. Guo, Y. Q. Li, T. Yu, T. G. Zhang, J. N. Zhang, H. Liu, F. G. Liu, X. J. Xie, Q. Zhu, and Y. A. Zhao, "Diagnosis of gastric intestinal metaplasia with confocal laser endomicroscopy in vivo: a prospective study," *Endoscopy* **40**(7), 547–553 (2008).
10. R. Kiesslich, G. K. Anagnostopoulos, A. Axon, K. Deinert, H. Neuhaus, M. Goetz, E. Coron, O. Pech, M. Vieth, C. Schneider, P. R. Galle, and M. Neurath, "Interobserver variation and standardized training for confocal laser endomicroscopy image interpretation in the upper and lower GI tract," *Gastrointest. Endosc.* **65**(5), AB354–AB354 (2007).
11. T. J. Muldoon, S. Anandasabapathy, D. Maru, and R. Richards-Kortum, "High-resolution imaging in Barrett's esophagus: a novel, low-cost endoscopic microscope," *Gastrointest. Endosc.* **68**(4), 737–744 (2008).
12. T. Muldoon, M. C. Pierce, D. Nida, M. D. Williams, A. Gillenwater, and R. Richards-Kortum, "Subcellular resolution molecular imaging within living tissue by fiber microendoscopy," *Opt. Express* **15**(25), 16413–16423 (2007).
13. F. Argenti, L. Alparone, and G. Benelli, "Fast algorithms for texture analysis using cooccurrence matrices," *IEE Proc. F, Radar Signal Process.* **137**(6), 443–448 (1990).
14. T. Collier, M. Guillaud, M. Follen, A. Malpica, and R. Richards-Kortum, "Real-time reflectance confocal microscopy: comparison of two-dimensional images and three-dimensional image stacks for detection of cervical precancer," *J. Biomed. Opt.* **12**(2), 024021 (2007).
15. K. W. Gossage, T. S. Tkaczyk, J. J. Rodriguez, and J. K. Barton, "Texture, analysis of optical coherence tomography images: feasibility for tissue classification," *J. Biomed. Opt.* **8**(3), 570–575 (2003).
16. S. Srivastava, J. J. Rodriguez, A. R. Rouse, M. A. Brewer, and A. F. Gmitro, "Computer-aided identification of ovarian cancer in confocal microendoscope images," *J. Biomed. Opt.* **13**(2), 024021 (2008).
17. C. A. Lingley-Papadopoulos, M. H. Loew, M. J. Manyak, and J. M. Zara, "Computer recognition of cancer in the urinary bladder using optical coherence tomography and texture analysis," *J. Biomed. Opt.* **13**(2), 024003 (2008).
18. L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Wadsworth International Group, Belmont, California (1984).
19. M. Conio, S. Bianchi, G. Lapertosa, R. Ferraris, R. Sablich, S. Marchi, V. D'Onofrio, T. Lacchin, G. Iaquinto, G. Missale, P. Ravelli, R. Cestari, G. Benedetti, G. Macri, R. Fiocca, F. Munizzi, and R. Filiberti, "Long-term endoscopic surveillance of patients with Barrett's esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study," *Am. J. Gastroenterol.* **98**(9), 1931–1939 (2003).
20. N. Morgan and H. Bourlard, "Generalization and parameter estimation in feedforward nets: some experiments," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, Ed., Vol. **2**, pp. 630–637, Morgan Kaufmann Publishers, Inc., San Mateo, CA (1990).
21. V. S. Wang, J. L. Hornick, J. A. Sepulveda, R. Mauer, and J. M. Ponomeros, "Low prevalence of submucosal invasive carcinoma at esophagectomy for high-grade dysplasia or intramucosal adenocarcinoma in Barrett's esophagus: a 20-year experience," *Gastrointest. Endosc.* **69**(4), 777–783 (2009).
22. R. Kiesslich, M. Goetz, K. Lammersdorf, C. Schneider, J. Burg, M. Stolte, M. Vieth, B. Nafe, P. R. Galle, and M. F. Neurath, "Chromoscopy-guided endomicroscopy increases the diagnostic yield of intraepithelial neoplasia in ulcerative colitis," *Gastroenterology* **132**(3), 874–882 (2007).
23. V. Becker, M. Vieth, M. Bajbouj, R. M. Schmid, and A. Meining, "Confocal laser scanning fluorescence microscopy for in vivo determination of microvessel density in Barrett's esophagus," *Endoscopy* **40**(11), 888–891 (2008).
24. W. L. Curvers, R. Singh, L. M. Song, H. C. Wolfsen, K. Ragnath, K. Wang, M. B. Wallace, P. Fockens, and J. J. Bergman, "Endoscopic tri-modal imaging for detection of early neoplasia in Barrett's esophagus: a multi-centre feasibility study using high-resolution endoscopy, autofluorescence imaging and narrow band imaging incorporated in one endoscopy system," *Gut* **57**(2), 167–172 (2008).