Dig Dis Sci (2011) 56:805–811
DOI 10.1007/s10620-010-1371-6

ORIGINAL ARTICLE

Quantitative Assessment of Endoscopic Images for Degree of Villous Atrophy in Celiac Disease

Edward J. Ciaccio · Govind Bhagat · Christina A. Tennyson · Suzanne K. Lewis · Lincoln Hernandez · Peter H. R. Green

Received: 10 June 2010 / Accepted: 26 July 2010 / Published online: 16 September 2010

Abstract

Background The degree of villous atrophy in celiac disease is difficult to assess at endoscopy. We sought to develop a quantitative technique for the evaluation of villous atrophy in endoscopic images.

Method In ten celiac patients as identified by standard endoscopy with biopsy, and ten control patients, standard and videocapsule endoscopic images of the duodenum were digitized. Subimages 7.5×7.5 mm^2 in area from random locations within each image were assessed by measuring the length of mucosal fissures per unit area (L), and correlating L with the histologic grade of villous atrophy as determined by modified Marsh criteria.

Results Mean L values for standard endoscopic images were 37.8, 43.3, 64.1, and 83.5 mm for Marsh grades II, IIIa, IIIb, and IIIc, respectively. Mean L values for videocapsule images were 49.1, 50.0, 64.7, and 72.4 mm for Marsh grades II, IIIa, IIIb, and IIIc, respectively. Significant differences in the means existed between celiac images (Marsh scores II-IIIc) versus controls ($p < 0.001$) for both endoscopic and videocapsule images. There were no significant differences between measurements obtained from endoscopic versus videocapsule images.

Conclusions Quantified image analysis correlates with the histologic grade of villous atrophy, is automated, and lacks observer bias, thus lending itself to standardization.

Keywords Celiac disease · Endoscopy · Videocapsule · Villous atrophy

E. J. Ciaccio · G. Bhagat · C. A. Tennyson · S. K. Lewis · L. Hernandez · P. H. R. Green
Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA

G. Bhagat
Department of Pathology, Columbia University College of Physicians and Surgeons, New York, NY, USA

E. J. Ciaccio (✉)
Celiac Disease Center at Columbia University, 180 Fort Washington Avenue, HP804, New York, NY 10032, USA
e-mail: ciaccio@columbia.edu

Introduction

The diagnosis of celiac disease is established by endoscopy and duodenal biopsy [1]. Studies have attempted to define the presence of villous atrophy at the time of endoscopy, mainly to determine the need for biopsy, and best candidate sites for biopsy. These techniques have included magnification endoscopy, with and without chromoendoscopy [2–5], and immersion techniques [6, 7]. In addition, with the increasing use of videocapsule endoscopy there is reliance on visual determination of endoscopic signs of villous atrophy for the diagnosis of celiac disease [8–11]. Application of specialized imaging techniques such as endoscopic confocal microscopy, a new diagnostic tool that enables subsurface microscopic imaging of live tissue during endoscopy, has now been applied to celiac disease diagnosis [12, 13]. However, use of this technique requires specialized equipment.

There is widespread use of routine video endoscopy (esophago-gastro-duodenoscopy, EGD) and videocapsule endoscopy (VCE). These techniques are however subjective in nature in that they require visual recognition of abnormal images. In order to overcome subjective observations we correlated image quantification techniques obtained by EGD and VCE with degree of villous atrophy assessed microscopically from endoscopic small bowel biopsies [14, 15].

Methods

Images from retrospectively identified patients who were undergoing endoscopy for the diagnosis of celiac disease, and from controls with normal appearing mucosa were used. The data were obtained from ten sequential celiac patients and ten controls. In each patient, non-magnified, color endoscopic images (527 × 444 pixels, Olympus EVIS EXERA II 180) and six biopsies were acquired at random from the descending duodenum. Biopsy specimens were obtained with standard forceps at approximately the same locations as were imaged. The biopsy specimens were fixed in formalin (10% neutral buffered formalin) and then embedded in paraffin, and the sections were stained with hematoxylin and eosin (H&E) stain. Images of biopsy slides were obtained at 400× magnification using a digital camera mounted on an Olympus bright field microscope and stored as high-resolution TIFF files for morphometric analysis. The slides were blindly reviewed and the degree of villous atrophy classified using the scoring system for villous atrophy developed by Marsh [14] and modified by Oberhuber [15]. The crypt-to-villous ratio was assessed in areas of the biopsies where at least three well-oriented villi could be identified.

Examples of endoscopic images that correlated with different degrees of villous atrophy are shown in Fig. 1a. Fissures were identified in each standard endoscopic image from the descending duodenum as abnormal deep furrows in the intestinal mucosa evident as crisscross lines with darker shading, and measured as shown in Fig. 1b. A 7.5 × 7.5 mm^2 area on the computerized grid, as estimated by the approximate image resolution, was selected at random and extracted to create a subimage which was used for edge detection and measurement [16]. A representative area used for analysis is delineated by the square in Fig. 1b. The subimage was extracted with ImageJ (Ver. 1.36b, National Institutes of Health, Bethesda, MD, a public domain Java image-processing program), which is a downloadable application for PC-type computers running a Java virtual machine. The ImageJ program is capable of displaying, editing, analyzing, and processing digital images. The subimage was identified in the ImageJ program by outlining a boxed area of the original image using the computer mouse. The boxed area was then automatically extracted by ImageJ and magnified for ease of visualization. The magnified subimage was skeletonized, which is a process of reducing an image to its line, or skeletal, structure while preserving the extent and connectivity of the original image patterns [17, 18]. Thus, by skeletonization, mucosal fissures appearing in the digital image were observed as/represented by solid lines. An automated routine was then used to compute the total length L of the skeletonized features within the subimage, where L is a real number and is expressed as a function of mucosal surface area. The time required for semiautomated measurement of L in each subimage was less than 60 s. In subsequent analyses the L measurement algorithm will be made fully automated, so that the expected time for delineation of each subimage will be less than 1 s.

The degree of villous atrophy was then assessed from the biopsies and classified using the modified Marsh criteria [14, 15]. The L value as calculated from endoscopy images was compared with the degree of villous atrophy. Images were also obtained by videocapsule endoscopy

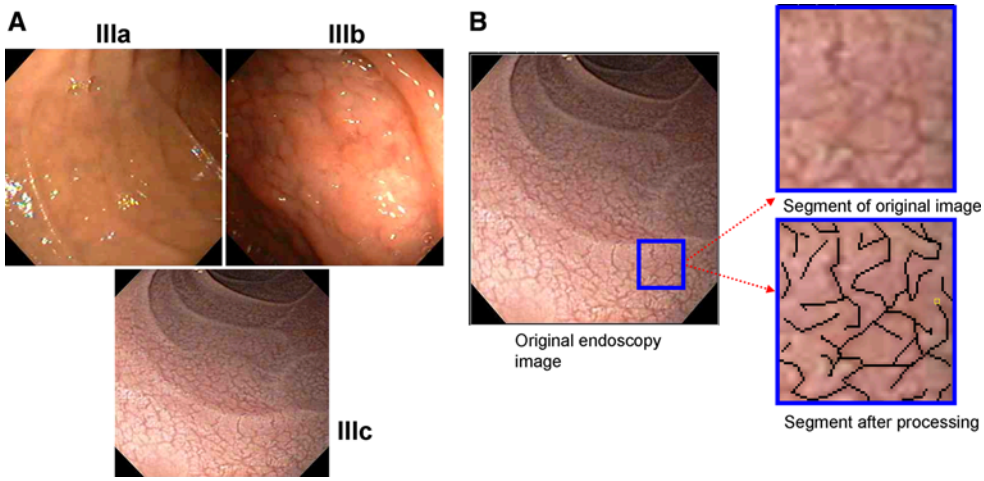

Fig. 1 Example of standard endoscopic images and subimage selection. **a** Images with Marsh scores of IIIa (partial villous atrophy), IIIb (subtotal villous atrophy), and IIIc (total villous atrophy). Increased abnormality of the mucosal surface is evident as the degree of villous atrophy increases. **b** Extraction of subimage for quantitative analysis. The total length L of the fissured pattern (*curved black lines*) per united area is calculated for quantitative analysis

from the same patients with celiac disease ($n = 10$) and the same control patients ($n = 10$). The Given® Diagnostic System with PillCam TM SB2 Capsule (576 × 576 pixels, Given Imaging Ltd, Yoqneam, Israel) was used to obtain videocapsule images, which were acquired from the descending duodenum as in standard endoscopy. Quantitation of the videocapsule images was done in the same manner as for standard endoscopy as described above. Biopsies had been performed within 3 months of videocapsule endoscopy.

In order to demonstrate the sensitivity of this method, we performed similar analysis on images from different celiac and control patients ($n = 10$ for each), using 2–8 subimages from each patient for analysis. A total of 51 subimages from standard endoscopy and 56 subimages from videocapsule endoscopy were analyzed from the 20 patients (107 in total). Subimage extraction was limited to image areas where the surface plane of the small intestinal mucosal surface was approximately perpendicular to the videocamera plane. For each extracted subimage, the mean ± SD of the L value was calculated. One-way ANOVA and the unpaired *t* test were used to determine the statistical significance of the difference between the means of the L value when binned according to the histopathologic grade determined from biopsy assessment ($p < 0.05$). The graphs and statistical analyses were generated using SigmaPlot (Systat Software ver. 9.0, 2004) and MedCalc (MedCalc Statistical Software ver. 9.5, 2008).

Results

Standard Endoscopy Results

Endoscopic images that correlated with the different degrees of villous atrophy were obtained and subimages were extracted from them, examples of which are shown in Fig. 2. To the left are original subimages and to the right are the same subimages following skeletonization. For images from patients with normal biopsies, the total fissure length L (in millimeters) ranged from 2 to 7 mm in the figure. For celiac patients, the value for L varied according to the Marsh grade. In Marsh grade II-IIIa lesions, the L value ranged from 35 to 43 mm, for Marsh IIIb it ranged from 60 to 68 mm (Fig. 2), and for Marsh IIIc the range was 74–96 mm (not shown). In Fig. 3, the standard endoscopy data is depicted as a scatterplot, with means shown as straight lines through the points, and quantitative values given in Table 1. Each point represents a subimage measurement, and they mostly cluster in one location for each Marsh type and the controls. The celiac patient data is shown in the columns labeled Marsh score II–IIIc while the control data is shown as a separate column in the graph. There is a clear and evident decrease in L score in correspondence with decreasing Marsh score. The variation in L score for any particular Marsh class is noted as a standard deviation in Table 1, and these values are substantially smaller than the means (i.e., coefficients of variation are

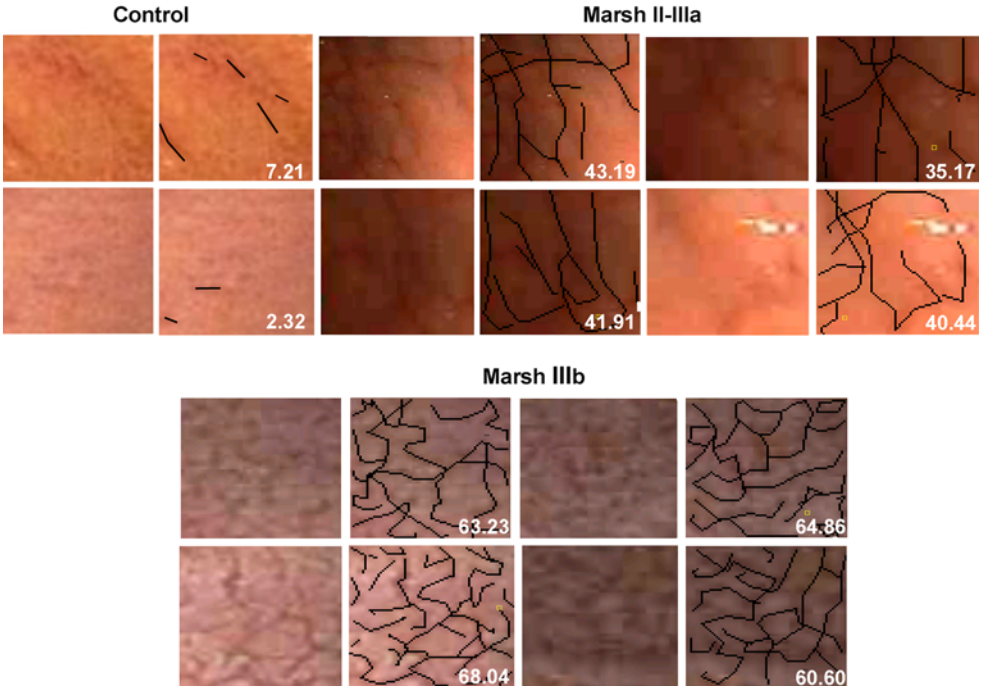

Fig. 2 Analysis of endoscopic subimages. An example result for the measurement of L value in standard endoscopy imaging is shown. In each panel the fissures are detected and the L value is calculated. L values increases from normals to celiacs with a grade of Marsh II or IIIa, to Marsh IIIb or IIIc (not shown)

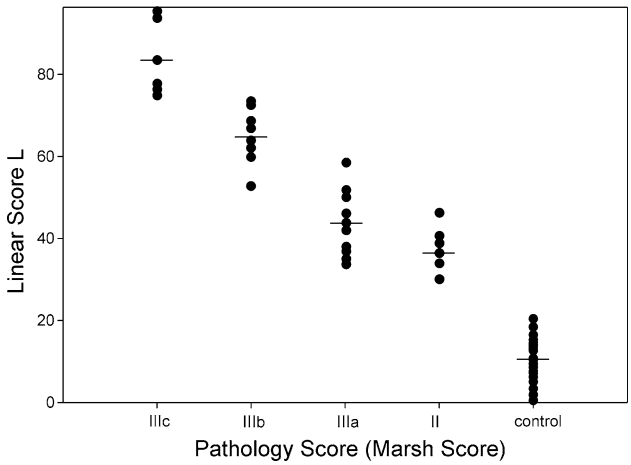

Fig. 3 Standard endoscopy assessment. Scatterplot of linear score L value versus integer histologic grade in standard endoscopy imaging (each point represents measurement from one subimage). There is a significant difference between the means based on one-way ANOVA ($p < 0.001$)

Table 1 Summary values of linear score L

Marsh/type	Standard endoscopy	Videocapsule endoscopy	No. of patients
IIIc	83.53 ± 9.51	72.42 ± 9.99	3
IIIb	64.07 ± 7.12	64.69 ± 4.01	1
IIIa	43.32 ± 7.35	49.96 ± 9.92	4
II	37.76 ± 5.61	49.01 ± 14.45	2
Cont	10.13 ± 4.97	11.43 ± 5.21	10

Table 1 shows summarized values in millimeters, which are graphed in Figs. 3 and 5

Endoscopy versus videocapsule L value pairs are not significantly different

MN, mean value; *SD*, standard deviation; *cont*, controls

small), suggesting that the method has good reproducibility. The mean difference between the values from the patients with celiac disease and controls was significant ($p < 0.001$, Table 2). There was also a significant difference between the mean values for Marsh IIIc versus Marsh IIIb, and between Marsh IIIb versus Marsh IIIa images (Table 2). This suggests that there is an underlying pathology can be related to image features, specifically the fissures that can be delineated.

Videocapsule Endoscopy Results

We then applied this technique to images obtained by VCE as shown in Fig. 4 for a control patient (left) and a celiac disease patient with Marsh IIIa pathology (right). The endoscopy images for Marsh IIIa exhibit scalloping of the edges of the mucosal folds and some fissuring. The control images exhibit normal folds and minimal fissuring. Examples of subimages taken at random for analysis of these images and the L values are shown. Significant fissuring is detected in the Marsh IIIa subimages. Notice that subimages are extracted from areas of the image in which the videocamera is approximately perpendicular to the orientation of the surface plane. This is a requirement of the method; thus no subimages were extracted which would extend between luminal folds. In the control subimages shown there is almost no fissuring detected. The videocapsule assessment is summarized for areas classified by Marsh score (celiac patients) and for controls in Fig. 5. Mean values of L are shown as horizontal lines and are stated quantitatively in Table 1 with standard deviations given. As for standard endoscopy image results, the standard deviations are substantially smaller than the means, suggesting that the measurements are reproducible. The mean L value for any of the Marsh scores (celiac patients) was significantly different from controls ($p < 0.001$, Table 2). There was also a significant difference between the mean values for Marsh IIIb versus Marsh IIIa videoendoscopy images (Table 2).

Correlation of Histopathology with Endoscopy

For both endoscopic and videocapsule image sets, mean L value increased in the order Marsh II < Marsh IIIa < Marsh IIIb < Marsh IIIc suggesting that there is a direct relationship between the degree of villous atrophy and the amount of fissuring that is detectable as measured by L. The statistical characteristics of the images measured by L are thus reflective of the degree of injury as classified by the Marsh score, and generalizable to different imaging

Table 2 Significance of mean L value between marsh scores

Type/marsh	IIIc vs. IIIb	IIIb vs. IIIa	IIIa vs. II	IIIc, IIIb, IIIa, II vs. cont
Endoscopy	$p < 0.001$	$p < 0.001$	NS	$p < 0.001$
Videocapsule	NS	$p < 0.02$	NS	$p < 0.001$

Table 2 shows significance between measurement sets

Set 1, significance between Marsh IIIc and Marsh IIIb; *Set 2*, significance between Marsh IIIb and Marsh IIIa; *Set 3*, significance between Marsh IIIa and Marsh II; *Set 4*, significance between Marsh II and controls; *NS*, not significant; *cont*, controls

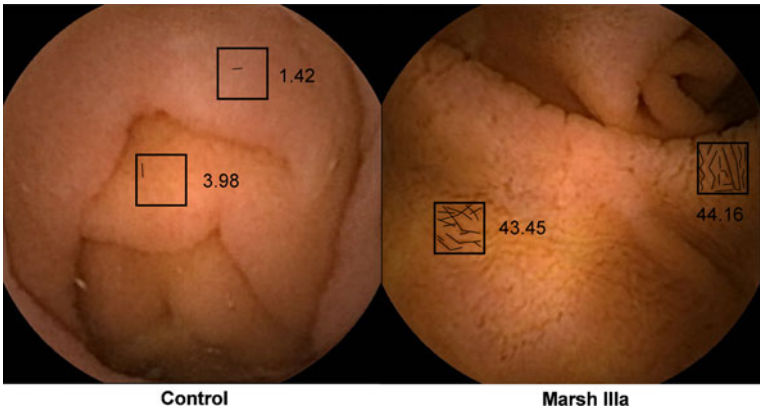

Fig. 4 Analysis of videocapsule images. Examples of videocapsule images from the descending duodenum for a control patient and celiac patient are shown. Subimages and measured L values are given. For each patient, measurements were made from two or more subimages and from different images. Substantial fissuring is evident in the Marsh IIIa image as well as scalloping of mucosal folds

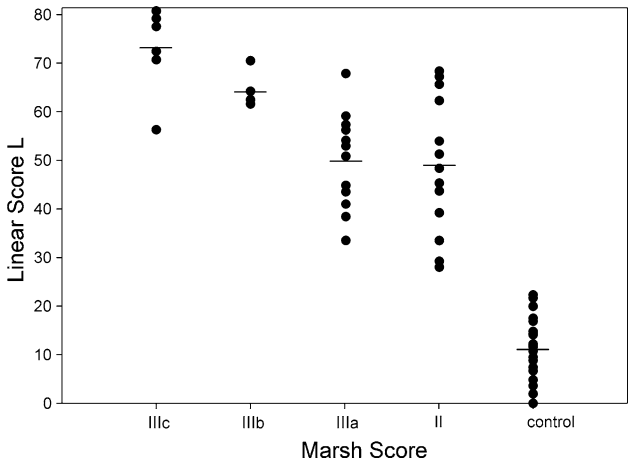

Fig. 5 Videocapsule endoscopy assessment. Scatterplot of linear score L value versus integer histologic grade in videocapsule endoscopy imaging. Each point represents measurement from one subimage. There is a significant difference between the means ($p < 0.001$)

modalities (standard endoscopy and videocapsule endoscopy in this study).

Discussion

We used a quantitative image-processing technique to grade the severity of the abnormalities of standard and videocapsule endoscopic images of duodenal mucosa of patients with celiac disease. This study shows that the degree of villous atrophy can be quantified using a mathematical method and endoscopic images, similar to what was previously demonstrated for digitized images of small-bowel biopsies [19]. Flat surface areas perpendicular to the videocamera angle and away from luminal folds, where fissures would be readily detectable, were specifically targeted to measure L. Yet, this technique could also be applied to other disease types, either as is (in which case it would be expected to measure different values depending on the characteristics) or by altering the algorithm to specifically target the image statistical characteristics of other types of lesions and pathological conditions. Thus, it would be possible to detect and measure varied injury patterns such as those present in autoimmune enteritis, inflammatory bowel disease, tropical sprue, and infectious enteritidis.

The findings of this study show that in celiac disease, L decreases with Marsh score in both standard and videocapsule endoscopy, as shown in Figs. 3 and 5, respectively. Since image processing is used for quantification, these measures are objective and not subject to interobserver variability. Moreover, they provide real-numbered scores, which potentially reflect the state of the intestinal mucosa more precisely and may be useful for detection of low-grade villous atrophy that would not be readily apparent by visual inspection alone. Additionally, the entire quantification process can be automated for more rapid and accurate assessment of villous atrophy in patients suspected of having celiac disease or another disorder associated with mucosal architectural abnormalities, subject of future work. Although only images from the descending duodenum were quantified in this study, we believe that the measurement of L can be extended to other areas of the small intestine so long as biopsies can be obtained for initial verification, the subject of future work. The algorithm can potentially be developed for real-time analysis using a dedicated computer accessing the standard or videocapsule imaging process. In addition, comparisons using this automated, objective technique, could be made both between sequential studies of individual patients and between different patients.

Our study suggests that endoscopic image manipulation may be a surrogate for analysis of biopsies, especially applicable to videocapsule endoscopy. In this preliminary study, we used images from patients with celiac disease, all of whom had visual abnormalities detected by standard, non-magnified endoscopic images, videocapsule images, and corroborated by histologic review of duodenal biopsies that displayed varying grades of villous atrophy. No

patients with Marsh I lesions, characterized by complete absence of villous atrophy, were studied. There were however two patients with Marsh II grade abnormalities.

Villous atrophy is present in 85% of patients with celiac disease, with total villous atrophy (Marsh IIIc) being significantly more frequent in the distal duodenum [20], the area from which images were obtained for this study. Approximately half of all celiac patients have identical degrees of villous atrophy throughout the duodenum, with no areas that are histologically normal [20]. Uncommonly, the duodenum can be entirely normal by standard endoscopic examination while the proximal and distal small intestine show classic features of celiac disease by videocapsule [21]. Yet villous atrophy may be present even at such areas of normal endoscopic appearance [22], and is potentially revealed by image magnification, an approach that could be subjected to our quantitative technique.

Although the brightness of images obtained at endoscopy and videocapsule endoscopy was uneven (see Fig. 1a), the process can be automated to correct for uneven lighting. A two-dimensional curve fit of the mean brightness level can be subtracted from the image [17] so that the corrected image will then have uniform brightness.

Quantitative image processing applied to capsule endoscopy can potentially be useful to assess the patchiness of disease [23–26] and improvement in mucosal appearance after institution of a gluten-free diet [10], a possible future application of our method. To validate the usefulness of our new method in cases of patchy lesions will require acquisition of endoscopic biopsies at each subimage location where L is measured. The technique may be useful for detection of very low-grade villous atrophy in videoclip images [27] as well as to detect patchiness. Yet, whether quantitative image analysis is sensitive enough to pick up this type of lesion will require further study. New imaging modalities such as magnification chromoendoscopy may be better suited for noninvasive study of low-grade lesions. The quantification of mucosal abnormalities by objective techniques could also be used for the assessment of patients with celiac disease enrolled in drug trials, where investigators and study subjects would prefer noninvasive procedures. Additionally, this technique may be applicable for study of other diseases with small bowel lesions including inflammatory bowel disease, but will need to be adapted to the statistical characteristics of the lesions.

Limitations

This study only included patients with celiac disease who had an abnormal appearing small bowel mucosa and villous atrophy on biopsy, and a set of control patients with normal mucosal appearance. It did not include celiac patients with increase of intra-epithelial lymphocytes that lacked villous atrophy, the subject of future study. Subimage measurements were not necessarily obtained from the same exact location as the biopsy. Furthermore, an assessment of different segments of the duodenum, for example the duodenal bulb, and comparison between these and biopsy classification, was not done presently but is planned for future work. Due to the patchiness of the disease, more extensive mapping at multiple different sites in the small bowel will be required to address this issue. The surface area of extracted subimages was estimated based upon approximate luminal dimensions and camera-mucosa distance. A more exact estimate of surface area would require calibrated surface dimensions in the image field-of-view. Image resolution is currently somewhat limited but can potentially be improved without increasing the sample size by using compression algorithms [28]. Were a high definition endoscope or narrow band imager to be utilized, it could potentially improve the power of the algorithm to distinguish normal from abnormal villi dramatically, by enhancing the detail in the image structure, the subject of future work. Another issue in this study is the fact that we are imaging constrained areas of the mucosa, while comparisons are made with biopsies that were assumed to be representative of the whole area.

Acknowledgments The implementation of this study was made possible in part with a grant from the Celiac Sprue Association (CSA) Peer Review Research Grant Program.

References

1. Green PH, Cellier C. Celiac disease. *N Engl J Med*. 2007;357:1731–1743.
2. Hurlstone DP, Sanders DS. High-magnification immersion chromoscopic duodenoscopy permits visualization of patchy atrophy in celiac disease: an opportunity to target biopsies of abnormal mucosa. *Gastrointest Endosc*. 2003;58:815–816.
3. Siegel LM, Stevens PD, Lightdale CJ, et al. Combined magnification endoscopy with chromoendoscopy in the evaluation of patients with suspected malabsorption. *Gastrointest Endosc*. 1997;46:226–230.
4. Lo A, Guelrud M, Essenfeld H, Bonis P. Classification of villous atrophy with enhanced magnification endoscopy in patients with celiac disease and tropical sprue. *Gastrointest Endosc*. 2007;66:377–382.
5. Banerjee R, Shekharan A, Ramji C, et al. Role of magnification endoscopy in the diagnosis and evaluation of suspected celiac disease: correlation with histology. *Indian J Gastroenterol*. 2007;26:67–69.
6. Gasbarrini A, Ojetti V, Cuoco L, et al. Lack of endoscopic visualization of intestinal villi with the "immersion technique" in overt atrophic celiac disease. *Gastrointest Endosc*. 2003;57:348–351.
7. Cammarota G, Cesaro P, La Mura R, et al. Role of the "immersion technique" in diagnosing celiac disease with villous

atrophy limited to the duodenal bulb. *J Clin Gastroenterol*. 2007;41:571–575.
8. Cellier C, Green PH, Collin P, Murray J. ICCE. ICCE consensus for celiac disease. *Endoscopy*. 2005;37:1055–1059.
9. Hopper AD, Sidhu R, Hurlstone DP, McAlindon ME, Sanders DS. Capsule endoscopy: an alternative to duodenal biopsy for the recognition of villous atrophy in coeliac disease? *Dig Liver Dis*. 2007;39:140–145.
10. Murray JA, Rubio-Tapia A, Van Dyke CT, et al. Mucosal atrophy in celiac disease: extent of involvement, correlation with clinical presentation, and response to treatment. *Clin Gastroenterol Hepatol*. 2008;6:186–193.
11. Biagi F, Rondonotti E, Campanella J, et al. Videocapsule endoscopy and histology for small-bowel mucosa evaluation: a comparison performed by blinded observers. *Clin Gastroenterol Hepatol.* 2006.
12. Trovato C, Sonzogni A, Ravizza D, et al. Celiac disease: in vivo diagnosis by confocal endomicroscopy. *Gastrointest Endosc*. 2007;65:1096–1099.
13. Leong RW, Nguyen NQ, Meredith CG, et al. In vivo confocal endomicroscopy in the diagnosis and evaluation of celiac disease. *Gastroenterology*. 2008;135:1870–1876.
14. Marsh MN. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology*. 1992;102:330–354.
15. Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease: time for a standardized report scheme for pathologists. *Eur J Gastroenterol Hepatol*. 1999;11:1185–1194.
16. Angelini ED, Ciaccio EJ. Optimized region finding and edge detection of knee cartilage surfaces from magnetic resonance images. *Ann Biomed Eng*. 2003;31:336–345.
17. Russ JC. *The Image Processing Handbook*. 5th ed. Boca Raton: CRC Press; 2007, pp 498–510.
18. Ciaccio EJ, Costeas C, Coromilas J, Wit AL. Static relationship of cycle length to reentrant circuit geometry. *Circulation*. 2001;104: 1946–1951.
19. Ciaccio EJ, Bhagat G, Naiyer AJ, Hernandez L, Green PH. Quantitative assessment of the degree of villous atrophy in patients with coeliac disease. *J Clin Pathol*. 2008;61:1089–1093.
20. Ravelli A, Bolognini S, Gambarotti M, Villanacci V. Variability of histologic lesions in relation to biopsy site in gluten-sensitive enteropathy. *Am J Gastroenterol*. 2005;100:177–185.
21. Muhammad A, Pitchumoni CS. Newly detected celiac disease by wireless capsule endoscopy in older adults with iron deficiency anemia. *J Clinical Gastroenterology*. 2008;42:980–983.
22. Oxentenko AS, Grisolano SW, Murray JA, Burgart LJ, Dierkhising RA, Alexander JA. The insensitivity of endoscopic markers in celiac disease. *Am J Gastroenterol*. 2002;97:933–938.
23. Wadge E, Boulougoura M, Kodogiannis V. Computer-assisted diagnosis of wireless-capsule endoscopic images using neural network based techniques. CIMSA 2005—IEEE International Conference on Computational Intelligence for Measurement Systems and Applications Giardini Naxos, Italy, July 20–22, 2005:328–333.
24. Kodogiannis V, Boulougoura M. Neural network-based approach for the classification of wireless-capsule endoscopic images. IEEE International Joint Conference on Neural Networks (IJCNN). IEEE Computer Society, USA, 2005:2423–2428.
25. Mackiewicz M, Berens J, Fisher M. Wireless capsule endoscopy color video segmentation. *IEEE Trans Med Imaging*. 2008;27: 1769–1781.
26. Vecsei A, Fuhrmann T, Liedlgruber M, Brunauer L, Payer H, Uhl A. Automated classification of duodenal imagery in celiac disease using evolved Fourier feature vectors. *Comput Methods Programs Biomed.* 2009;95(2 Suppl):S68–S78.
27. Ciaccio EJ, Tennyson CA, Lewis SK, Krishnareddy S, Bhagat G, Green PH. Distinguishing patients with celiac disease by quantitative analysis of videocapsule endoscopy images. *Comput Methods Programs Biomed*. 2010;100:39–48.
28. Lin MC, Dung LR, Weng PK. An ultra-low-power image compressor for capsule endoscope. *Biomed Eng Online*. 2006;5:14/1–8.