Real-time reflectance confocal microscopy: comparison of two-dimensional images and three-dimensional image stacks for detection of cervical precancer

1 Introduction

Over the last 50 y, the overall age-adjusted mortality rate due to cancer has not changed in the United States.1 It is well recognized that precancerous lesions precede the development of many epithelial cancers; recognizing and treating these lesions can reduce both the incidence and mortality of cancer. Currently, detection of precancerous lesions in organs such as the uterine cervix, skin, oral mucosa, colon, and bladder relies on visual examination of the tissue at risk. Biopsies of areas that appear suspicious are sectioned, stained, and examined by pathologists to confirm the diagnosis of precancer. This process is invasive, labor intensive, and expensive. As a result, the opportunity to intervene and treat the disease when it is...
most responsive to therapy is often missed or substantially delayed. Better early detection tools to identify precancerous lesions \textit{in situ} are necessary.

Precancers of the epithelium are associated with a variety of morphological changes including increased nuclear size, pleomorphism, nuclear hyperchromasia, and increased nuclear to cytoplasmic ratio. Clinical pathologists currently assess these features qualitatively, by examining stained sections at low- and high-power magnification. Recently, a number of studies have shown that quantitative analysis of digital images of stained histologic sections can aid in the identification of precancers. In quantitative pathology, measurements of morphological and architectural features are used to make the diagnosis. In both cases, the morphologic information is acquired from stained tissue sections taken from biopsies, which are invasive, expensive, and painful and limit the area at risk that can be examined.

An alternative that does not require tissue removal, sectioning, and staining is \textit{in vivo} confocal microscopy. Confocal microscopy can provide images of intact tissue with subcellular resolution and has been proposed as a method for acquiring cell morphometry and epithelial architecture \textit{in vivo} using reflected light from the tissue. Reflected light is isolated from out-of-focus background light simply by placing a pinhole at the conjugate image plane. Variations in the index of refraction within cells provide the source of contrast to image cellular and nuclear morphometry. Confocal microscopy has been demonstrated to detect changes associated with precancerous and cancerous lesions in the skin, oral cavity, cervix, esophagus, and colon.

We previously demonstrated that nuclear morphometry acquired from two-dimensional (2-D) confocal images of the cervical epithelium can be used to distinguish normal epithelium from high-grade dysplasia in the cervix. In \textit{in vivo} confocal images are generally acquired with the image plane parallel to the epithelial surface, in contrast to transverse histologic sections, which are oriented perpendicular to the epithelial surface. However, a series of 2-D confocal images can be acquired at increasing depth beneath the epithelial surface to create a three-dimensional (3-D) stack of confocal images, thus enabling 3-D morphometry. The utility of 3-D image stacks of epithelial tissue for detecting dysplasia has not yet been explored. The goal of the study reported here was to extract morphometry and architecture information from 2-D confocal reflectance images and 3-D image stacks from fresh, unstained cervical biopsies and compare their potential for detecting dysplasia.

2 Materials and Methods

2.1 Patients

Cervical biopsies were acquired from 29 patients at the Colposcopy Clinic at the University of Texas M.D. Anderson Cancer Center, in Houston, Texas. The patients were referred to the clinic for suspected dysplasia on the basis of an abnormal cervical cytology or for removal of cervical tissue using the loop electrical excision procedure (LEEP) due to a previous diagnosis of dysplasia. Informed consent was obtained from each patient, and the project was reviewed and approved by the Surveillance Committee at the University of Texas M.D. Anderson Cancer Center and the Institutional Review Board at the University of Texas at Austin.

2.2 Specimens and Confocal Imaging

Four uterine cervix biopsies were acquired from each patient under colposcopic guidance. Two adjacent biopsies were taken from a colposcopically normal area, and two adjacent biopsies were taken from a colposcopically abnormal area of the cervix. Colposcopic impression (normal/abnormal) was recorded for each biopsy. One normal and abnormal biopsy pair was submitted for routine histologic examination using the hematoxylin and eosin (H&E) stain. Sections were examined by an experienced, board-certified gynecologic pathologist (AM), and classified as negative for dysplasia or cervical intraepithelial neoplasia (CIN) 1, 2, or 3 using standard diagnostic criteria. The other normal and abnormal biopsy pairs were placed in growth medium and reflectance confocal images were obtained within 4 h of excision.

Reflectance confocal images were obtained from each biopsy using an epillumination laser scanning confocal microscope described in Ref. 8, with minor modification. The illumination source was a continuous wave diode laser operating at 810 nm. Images were acquired with a 25 × 0.8 numerical aperture (NA) water immersion objective (Plan-Neofluar multimmersion, Zeiss). The measured lateral and axial resolution of the system are 0.8 and 2 to 3 μm, respectively. The beamsplitter used in Ref. 8 was replaced with a polarizing beamsplitter to increase the total system throughput. A quarter-wave plate was added between the beamsplitter and sample to rotate the polarization of the reflected signal light.

Immediately prior to imaging, biopsies were removed from the growth medium and rinsed with phosphate-buffered saline (PBS). Image frames were acquired at multiple epithelial depths in 1.5-μm intervals up to the 250-μm working distance of the microscope. Then, a 6% solution of acetic acid was added to each sample, and the image sequence was repeated. Images were digitized using a video frame grabber card and displayed at 6.5 frames/s on a computer monitor. Individual bitmap image files were saved from the frame grabber’s video buffer. Uncompressed video files were also acquired from the video buffer as each sample was moved toward the objective to acquire 3-D image stacks. The image depth was adjusted by moving the sample stage using a stepper motor.

2.3 Image Processing

Post-acetic-acid images from each biopsy were processed to segment the cell nuclei at each image depth; acetic acid substantially enhances nuclear backscattering. We used a segmentation algorithm to automatically segment cell nuclei within each image. The segmentation algorithm correctly identifies 90% of nuclei present, with a false positive error rate of 14% compared to hand segmentation. The algorithm is described in detail in Ref. 12. First, a nonlinear edge-preserving image filtering technique—anisotropic median diffusion—was used to increase the signal-to-background ratio. Next, knowing that the images of the nuclei exhibit strong local dependencies, while exhibiting a wide range of gray scales, we utilized a Gaussian Markov random field (GMRF) to model these dependencies independent of the overall gray-
scale variations. Then we reduced the number of nonnuclei by removing objects that are not present in successive frames and further reduced the number of nonnuclei using Bayesian classification. As a final step, minor corrections in the segmentation were performed by hand at each image depth to remove false positive nuclei based on visual analysis of images as the gold standard. No correction of the nuclear shape was performed. After segmentation, 2-D images taken at 50 μm below the tissue surface, and the entire 3-D image stacks were analyzed to extract several morphologic and architectural features. We selected the image obtained 50 μm beneath the surface for two reasons. First, previous studies have shown that confocal images obtained at this depth provide good separation between normal and high-grade cervical precancers.8 Second, studies to identify biomarkers of high-grade cervical precancer indicate that differences in biomarker expression differ most in the intermediate layer of tissue, located in the upper 2/3 of the epithelium.13 To narrow the number of features to include in the study, morphologic and architectural features that have previously been shown to aid in distinguishing between normal and dysplastic tissue were calculated, including features related to nuclear size, density, and nuclear shape.3 Previous studies indicate that nuclear backscattering correlates with nuclear size, nuclear density, chromatin content, and chromatin texture.14

The mean and standard deviation of the following morphologic features were extracted from both the 2-D and 3-D data sets: nuclear area (from 2-D images), nuclear volume (from 3-D image stacks), and maximum nuclear and nuclear elongation. For the 3-D stacks, the mean and standard deviation of each feature were calculated over the entire volume, excluding the layer of cells immediately above the basement membrane, while the mean and standard deviation of each feature were calculated from the single 2-D image acquired 50 μm beneath the surface.

The mean and standard deviation of the following architectural features were extracted: cell density, area disorder, volume disorder, mean distance between each cell and its neighbors (mean separation), and the minimum distance between each cell and its nearest neighbor (minimum separation). The architectural analysis of the images was based on the calculated Euclidian centroids of all detected nuclei from the segmentation algorithm. The neighborhood of each cell was determined by tessellation of the nuclear centroids using the algorithms in MATLAB, version 6.1. All cells with common boundaries were considered neighbors. Again, for the 3-D stacks, the mean and standard deviation of each feature were calculated over the entire volume, excluding the layer of cells immediately above the basement membrane, while the mean and standard deviation of each feature were calculated from the single 2-D image acquired 50 μm beneath the surface.

The epithelial light scattering coefficient was also extracted from each biopsy using the techniques described in Ref. 15. Briefly, the intensity of reflected light was measured as a function of depth from the 3-D image stacks within each biopsy. The depth-dependent reflectance intensity was fit to an exponential attenuation function to extract the attenuation coefficient assuming that light attenuation can be described using Beer’s law. In the near-IR, where confocal measurements are frequently made, the absorption coefficient is significantly less than the scattering coefficient, so that this process can be used to estimate the epithelial tissue scattering coefficient.16

The morphologic and architectural features extracted from the 2-D and 3-D data sets corresponding to biopsies diagnosed as histologically normal and CIN 2/3 were compared. A Student’s t test was used to test the hypothesis that the mean values of these parameters were different between normal and CIN 2/3. Combinations of morphologic features, and architectural features, as well as scattering coefficients, were analyzed to determine the diagnostic potential for discriminating CIN 2/3 from normal tissue.

3 Results

Fifty-eight biopsies were obtained from 29 patients participating in the study. Imaging and successful segmentation were performed on nine biopsies; five of these were diagnosed as squamous normal tissue and four were diagnosed as CIN 2 or CIN 2/3. The extensive tissue processing required for this study, which required 2-D and 3-D imaging, was not successful in 27 biopsies due to the small size or fragile nature of the biopsies. Images were not available from two biopsies due to a failure with the imaging software. Due to poor image contrast, automated image segmentation was not successful on images from the entire 3-D stack of 16 biopsies; although hand segmentation was possible in images of these biopsies it was not performed. Last, four biopsies were excluded due to a diagnosis of focal dysplasia or koilocytosis. For the biopsies with focal dysplasia, no meaningful correlation could be expected from these biopsies since the volume of dysplasia is significantly smaller than the volume imaged from each biopsy.

Figure 1 shows sample images taken from the stacks from a normal and a CIN 3 biopsy. Images were taken near the top of the epithelium, at an intermediate depth and near the basement membrane. In the upper epithelium of the normal biopsy, nuclei appear as bright disks sparsely dispersed within a grainy pattern, which is brightest at the periphery of cells. As the focal plane is moved toward the basement membrane, the nuclear density increases. In the image stack from the CIN 3 biopsy, the nuclei are large, bright, and densely packed throughout the entire epithelium. The signal decays rapidly as the focal plane of the microscope is moved down toward the basement membrane.

Figure 2 shows a 3-D representation of the size and location of epithelial nuclei within these biopsies after processing with the image segmentation algorithm. The spheres represent the location and size of the detected nuclei. Nuclear density increases with depth in the epithelium of the normal biopsy, while nuclear density is significantly higher throughout the epithelium of the CIN 3 biopsy.

We examined the morphologic and architectural features extracted from the 2-D images and the 3-D image stacks. One morphologic feature, the standard deviation of the maximum nuclear diameter, and three architectural features—the scattering coefficient, the mean cell separation, and the minimum cell separation—were found to provide significant separation between the normal and CIN 2/3 biopsies.

Figure 3 shows the mean of the standard deviation of the maximum nuclear diameter for all normal biopsies and all
CIN 2/3 biopsies; error bars indicate plus/minus one standard deviation. This feature was acquired by extracting the maximum diameter from each nucleus in each image set; the standard deviation of this value was calculated for each biopsy. Having obtained the standard deviation of the maximum nuclear diameter for each biopsy, we calculated the mean and variance of this parameter for all normal biopsies and for all CIN 2/3 biopsies. The standard deviation of the maximum nuclear diameter extracted from both the 2-D images and the 3-D image stacks shows a statistically significant difference between the normal and CIN 2/3 biopsies, with an increased separation in the feature extracted from the 3-D image stack ($p=0.003$).

Figure 4 compares the architectural feature that showed the most significant separation between normal tissue and CIN 2/3—the mean cell separation. This parameter was obtained by calculating the center-to-center distance of all epithelial cells in each biopsy. The mean value was calculated for each biopsy, and then these values were averaged for all normal and all CIN 2/3 biopsies. The mean cell separation is greater in normal tissue than in CIN 2/3 biopsies in all cases. The standard deviation of this parameter was substantially larger for the parameter extracted from the 2-D images than for the 3-D image stacks. The cell separation varies as a function of depth throughout the epithelium, decreasing from the surface to the basement membrane. While the 2-D images were all acquired 50 $\mu$m beneath the epithelial surface, the thickness of the epithelium varies from biopsy to biopsy. When the mean cell separation extracted from the 2-D images was scaled by the ratio of total epithelial thickness to the total epithelial thickness less 50 $\mu$m, the variance of this parameter was substantially reduced. The epithelial thickness was measured by subtracting the image depth at which the tissue surface and basement membrane were encountered. For biopsies where the basement membrane could not be visualized, the thickness was measured from the images of fresh transverse tissue slices. With this scaling, the standard deviation of the

Fig. 1 Confocal reflectance images of a normal (b) to (d) and a CIN 3 (f) to (h) cervical biopsy taken from a through-focus series; the focal plane is located 50 $\mu$m below the epithelial surface (b) and (f), at an intermediate layer (c) and (g), and near the basement membrane (d) to (h). A sample nuclear segmentation and tessellation (a) and (e) from the nuclei centers is illustrated above, and taken from, images (b) and (f), respectively.

Fig. 2 Three-dimensional representation of the location and volume of nuclei from two representative biopsies as determined from automatic segmentation of 3-D reflectance confocal image stacks for (A) a normal biopsy and (B) a CIN 3 biopsy.

Fig. 3 Comparison of the standard deviation of the maximum nuclear diameter as extracted from 2-D images and 3-D image stacks for normal biopsies and for CIN 2/3 biopsies.
cell separation in normal biopsies is substantially reduced. This scaled separation data are also presented in Fig. 4. This reduction in the standard deviation increases the significance of this 2-D feature, with the p-value decreasing from 0.035 to 0.004. The p value for the scaled 2-D parameter is less than that for the 3-D version of this feature.

We examined combinations of morphologic and architectural features to see which were most significantly different between the normal and CIN 2/3 groups. Specifically, we examined the products and ratios of all pairwise combinations of morphologic and architectural features. Figure 5(A) shows the ratio of the mean cell separation to the standard deviation of the maximum nuclear diameter; this feature showed the most significant differences between the normal and CIN 2/3 biopsies. Normal biopsies show greater mean cell separation and more uniform nuclear size; thus, this ratio is greater in the normal images as compared to the CIN 2/3 images. In this feature combination, the 2-D and 3-D features both show a significant separation with improved performance in the 3-D feature. Figure 5(B) shows the ratio of the total nuclear volume (or area in two dimensions) divided by the total cytoplasmic volume (or area in two dimensions), averaged over all normal and CIN 2/3 samples; this feature is frequently used by pathologists to detect the presence of CIN. In this case, only the 3-D feature produces a statistically significant separation between the normal and CIN 2/3 biopsies. This is likely because the 3-D stacks capture the flattening of the superficial nuclei, which is commonly seen in normal epithelium. The 2-D images obtained parallel to the epithelial surface do not capture this information.

Figure 6 compares the significance level (1-p value) of differences in the mean morphological features, architectural features, scattering coefficient, and feature combinations for normal and CIN 2/3 samples. Only features with p values less than 0.1 are shown. The smallest p values are obtained when the standard deviation of the maximum nuclear diameter is combined with the features that describe epithelial cell separation. Combinations using cell separation and nuclear size also performed well with the addition of the scattering information. In all cases where the 2-D features are multiplicatively scaled by the depth of acquisition divided by the epithelial thickness or scaled by the scattering coefficient, the significance level is equal to or greater than the comparable feature extracted from the 3-D image stacks. The 2-D features scaled by the epithelial scattering coefficient provide much more statistically significant differences than those scaled by the relative depth of image acquisition. Combinations of 3-D features scaled by scattering coefficient and relative acquisition depth are not shown since little change is seen in the significance level.

We examined whether these features could be used to discriminate between normal biopsies and CIN 2/3 biopsies. Figure 7(A) shows a scatter plot illustrating the mean N/C ratio extracted from the 2-D images and the scattering coefficient extracted from the 3-D image sets. Parameters extracted from the nine biopsies in this study are shown, along with data acquired from a previously published study of 19 cervical biopsies. The previously acquired data were used as training data to derive a linear discriminant function to separate normal tissue from CIN 2/3. A sensitivity and specificity of 100% is produced using this discriminant function. These data also suggest that the scattering coefficient provides slightly more discriminative capacity than the N/C ratio. Figures 7(B) and 7(C) show the values of the scattering coefficient and N/C ratio separately for each biopsy. The scattering coefficient alone provides perfect separation between the normal and CIN 2/3 biopsies.
Discussion

The mean quantitative features extracted from the 3-D confocal image stacks show more statistically significant separation between normal and CIN 2/3 biopsies than do the mean quantitative features captured from single 2-D images obtained 50 μm beneath the epithelial surface. However, when the 2-D image features are scaled by one additional piece of information about the epithelium, either the epithelial thickness or the epithelial scattering coefficient, then the 2-D features produce separation comparable to that extracted from 3-D image stacks. A limitation of this study is the relatively small sample size; further work with additional samples is required to confirm these conclusions and to investigate which scaled 2-D features provide greatest diagnostic capability. The limited sample size also prevented finding a distinction between normal and CIN 1 and between CIN 1 and CIN 2/3. However, it is encouraging that we find that using a discriminant function derived from previously published values for scattering coefficients from normal and CIN 2/3 biopsies, the scattering coefficients reported here were shown to discriminate between normal and CIN 2/3 with a sensitivity and specificity of 100%.

Note that the scattering coefficient appears to provide more discriminative capacity than the N/C ratio extracted from the confocal images. Previous work has shown a strong correlation between the epithelial scattering coefficient and increased DNA content and chromatin texture. These features are among those used by pathologists to assess whether a precancerous lesion is present. Increased scattering could also be a result of cytoplasmic scattering and changes in keratin and its precursors expressed in dysplastic epithelium.

The advantage of approaches that rely on scaled 2-D features is that much less image data is required; the 3-D image stacks require 100 to 300 times additional storage space and much less computational effort is required to extract features. Experimentally, using reflectance confocal imaging, it is much easier to acquire the epithelial scattering coefficient than the epithelial thickness. The scattering coefficient can easily be estimated using three depth-dependent confocal images acquired at 50 μm-micron intervals beneath the epithelial surface as previously reported, and does not require acquisition of images throughout the entire epithelial thickness. We found that acquiring the epithelial thickness by imaging through the epithelium is much more challenging. For example, in this study the increased scattering in the CIN 2/3 biopsies prevented the identification of the basement membrane in one of the four CIN 2/3 biopsies, primarily because the increased scattering coefficient limits imaging depth.

The separation achieved here compares well with the results of quantitative pathology in distinguishing between CIN 2/3 and normal tissue. In quantitative pathology, measurements of cellular and nuclear morphometric and architectural

![Fig. 6 Comparison of the level of significance (1-p values) for various individual features (A) and feature combinations (B) extracted from the 2-D images and the 3-D image stacks.](image)

![Fig. 7 Plots comparing the mean N/C ratio and epithelial scattering coefficient.](image)
features from histologic slides are used to attempt the diagnosis of pathology. Using a morphometric index calculated using multiple nuclear morphologic measurements of histology slides, MacAulay et al. reported a significant separation in the mean index score between normal and CIN 2/3 ($p=0.033$). The higher resolution and contrast from the stains enable more accurate segmentation and many more types of morphometric measurements than those made from the confocal images. The obvious advantage of the reflectance confocal microscopy system described here is its ability to image in vivo without the necessity of tissue removal, sectioning, or staining. The success of the scaled 2-D features has important implications for using this precancer detection method in the clinic. Acquisition of the epithelial thickness or scattering coefficient would require substantially less image acquisition and processing time than 3-D image sets.

References