Results of a pilot study of multispectral digital colposcopy for the in vivo detection of cervical intraepithelial neoplasia

Andrea Milbourne, Sun Young Park, J. Louis Benedet, Dianne Miller, Thomas Ehlen, Helen Rhodes, Anais Malpica, Jasenka Matisic, Dirk Van Niekerk, E. Neely Atkinson, Natalie Hadad, Nick MacKinnon, Calum MacAulay, Rebecca Richards-Kortum, Michele Follen*

Center for Biomedical Engineering, Department of Biostatistics and Applied Mathematics, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA
Department of Gynecologic Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA
Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78712, USA
Department of Optical Imaging and Gynecologic Oncology, The British Columbia Cancer Center, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA
Department of Gynecology, Obstetrics and Reproductive Sciences, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA

Available online 3 October 2005

Abstract

Objective. Fluorescence spectroscopy is a promising technology for the detection of cervical squamous intraepithelial precancers and cancers. To date, many investigators have focused on point spectroscopy as an adjunct to diagnostic colposcopy. A device that visualizes the whole field of the cervix is needed for screening. To that end, we have developed a multispectral digital colposcope that works through the colposcope to image with white light, UV excitation at 345 nm, and blue light at 440 nm excitation. Here, we report the pilot study that precedes a Phase I trial.

Methods. The MDC system is composed of a light source, a colposcope, and a video rate color CCD camera with a frame grabber and takes approximately less than 1 min to make images of the cervix. Patients were measured at baseline and after acetic acid placement with white light, 345 nm excitation, and 440 nm excitation from the xenon arc lamp. The white light is in the visible spectrum, 345 nm excitation in the UV spectrum and is not visible, and 440 nm excitation is blue light in the visible spectrum. White light generates a pink image of the cervix. 345 nm excitation, the UV light, excites fluorophores to emit a blue image. 440 nm excitation, the blue light, excites fluorophores to emit a green image. The patients underwent a loop excision procedure and the histopathology was inked and cut into 12 sections by the study pathologists. The histopathologic slides were scanned and the images were then reconstructed into maps. A diagnostic algorithm was calculated. The data were preprocessed, transformed, and analyzed by the K-means clustering method. Disease maps were generated using the algorithm and classifier and compared to white light colposcopy and the blue and green images obtained at 345 and 440 nm.

Results. Forty-six patients were measured at four clinical sites. Images were made of the cervix with white light, 345 nm excitation, and 440 nm excitation and are presented in the figures. As the study went on, images improved with improvements in the instrument. The white light and fluorescence images are presented with crudely constructed histopathologic maps and algorithmic maps. At 345 nm excitation, the UV light, histologically confirmed CIN appears darker blue; while at 440 nm excitation, the blue light, histologically confirmed CIN appears lighter green.

Conclusions. This pilot study shows that MDC images can be matched to both histopathologic and algorithmic maps. The device and the algorithm are evolving but show promise. A Phase I trial is planned.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Cervical cancer; Fluorescence spectroscopy; Colposcopy; Cervical intraepithelial neoplasia
Introduction

Fluorescence spectroscopy is a promising technology for the detection of precancerous lesions [1–3]. Point spectroscopy, using a fiber optic probe that touches tissue, has demonstrated a sensitivity and specificity in diagnosing cervical neoplasia of 86% and 74%, respectively, in diagnostic colposcopy clinics and 75% and 80%, respectively [4–7] in screening settings. This technology would best serve as an adjunct to diagnostic colposcopy, helping to better target biopsies. Differences in histologic diagnosis, menopausal status, and age among patients can cause variability in the measurements of fluorescence spectroscopy, while smoking, race, and menstrual cycle are not sources of variability [8–10]. Large screening and diagnostic trials are underway with the point probe [8–10].

In order to screen, one must see the whole field of the cervix. This type of field spectroscopy is being conducted in the lung, bladder, skin, and cervix [1,11–15]. We have developed a multispectral digital colposcope to visualize the cervix using white light, UV light at 345 nm, and blue light at 440 nm excitation. We chose these two wavelengths based on an analysis of the two most useful wavelengths for diagnosis of CIN from a large diagnostic trial of point spectroscopy [16].

The first step in the evaluation of a device is a pilot study. The objective of this report is to share preliminary but promising results of the pilot study.

Materials and methods

Instrumentation

In order to assess the entire field at risk, clinical application of this technology requires an imaging system. To meet this need, a multispectral digital colposcope (MDC) has been developed to measure multispectral images in a short period of time using autofluorescence and reflectance. The device has been described in detail [17]. The reflectance and the fluorescent light are captured and filtered using a commercially available video rate color charge-coupled device (CCD) camera costing less than $500. The spectrum of the measured light is resolved using the cyan, magenta, and yellow (CMY) channels of the camera.

The MDC system is composed of a light source, a colposcope, and a video rate color CCD camera with a frame grabber. A system diagram of the MDC system is depicted in Fig. 1(a). The light source is based on a xenon arc lamp, which provides monochromatic and broadband illumination. The fluorescence excitation light is produced using bandpass filters enclosed in a motorized filter wheel. Both 345-nm and 440-nm excitation wavelengths were chosen for the optimal excitation wavelengths. An attached fiber optic light source delivers the excitation light to the tissue. The MDC was built around a commercially available tilt-stand colposcope (Model IDL, Leisegang, Germany). This colposcope produces stereoscopic vision at 7.5× magnification. The original halogen lamp of the colposcope was removed to accommodate the fiber optic light guide. Fig. 1(b) shows a photograph of the MDC system. The camera is an analog video rate color CMY camera (CV S3200, JAI, Japan). The dynamic range of each channel is 8 bits. The camera CCD is sensitive to cyan, magenta, and yellow (CMY) colors; however, the camera transforms the measured CMY values into red, green, and blue (RGB) values. The operation of the camera was controlled using a PC computer and a LabVIEW program (National Instruments, Austin, TX). Fig. 2 reviews the spectrum of visible light.

Clinical measurements

The study protocol was reviewed and approved by the institutional review boards at The University of Texas M.D. Anderson Cancer Center, The University of Texas at Austin, The University of Texas Health Science Center at Houston, the Lyndon Baines Johnson and Harris County Health District, and the British Columbia Cancer Center and Cancer Research Center. Patients 18 years and older who were not pregnant and who were referred to the colposcopy clinic with an abnormal Papanicolaou smear were eligible. All patients signed an informed consent form and underwent a history, complete physical exam, Papanicolaou smear, cultures, pancolposcopy, and colposcopically directed biopsies. All the patients had high-grade lesions and were scheduled for treatment. All patients underwent measurement with the MDC and their scheduled loop electrosurgical procedure that day.

The MDC takes less than 1 min to make images of the cervix. Each patient was measured at baseline with white light, 345 nm excitation, and 440 nm excitation from the xenon arc lamp. Acetic acid (6%) was placed on the cervix for 2 min. A second image was taken with white light, 345 nm excitation and 440 nm excitation. Acetic acid was placed again briefly and the patient underwent colposcopy. The patients were then injected with lidocaine with epinephrine and underwent a loop electrosurgical excision procedure. The specimen was carried to the study pathologist for orientation and sectioning. The histopathology from the specimen was inked and cut by the study pathologists (AM, JM, DYN). Each ectocervix was cut into 12 pieces like a pie. The authors reviewed all the hematoxylin and eosin stained slides jointly. The images from the histopathologic slides were scanned and the shrinkage by formalin was corrected by image preprocessing. The images were then reconstructed using 3D Tool Software® by SYPr. Maps of the pathology were reviewed by the clinicians and pathologist and matched to the white light and MDC images.

The spectroscopic device for point spectroscopy is a non-significant risk device by U.S. Food and Drug Administration (FDA) standards [18]. The MDC is a non-significant risk device by IRB standards and the FDA application is pending. Since the device does not touch the patient, no disinfection is
necessary. Acetic acid enhances the differences in fluorescence between normal and dysplastic tissue [13].

**Calibration of device**

The MDC fluorescence imaging performance was tested by imaging different standard samples including solutions of three fluorescent compounds, Exalite (1.33 mM in ethylene glycol), flavin adenine dinucleotide (FAD) (0.016 mM in PBS) solution, and Rhodamine. The negative standard was the frosted cuvette. Fig. 3 shows the calibration standards for the device. The Exalite, FAD, and Rhodamine solutions were contained in 40–40 mm quartz cuvettes to cover the MDC’s entire field of view. This calibration assures us that the system is measuring accurately at 345 nm and 440 nm excitation [17]. The background signal for one image is plotted in Fig. 4. Here, we see that the red and green signal intensities are higher than the blue signal. Most of the signal for the blue channel is at the 0 value. Ideally, the histogram would have equal distributions of red, green, and blue signal. The camera and colposcope were realigned to produce a normal histogram.

**Data analysis**

Fluorescence intensity data were preprocessed at The University of Texas at Austin spectroscopy laboratory, M.D. Anderson’s Center for Biomedical Engineering, and at the British Columbia Research Center. Statistical analyses were performed using Matlab (Version 6.5, The MathWorks, Natick, MA). The fluorescence images acquired from the MDC were analyzed in three steps: image preprocessing, segmentation, and classification.

Image preprocessing was performed. First, median filtering was applied to each image to eliminate impulse noise spikes. Secondly, for segmentation, an aligning process
between white light and two fluorescence images was performed. In this step, an automatic image registration algorithm based on mutual information was developed to perform the image registration task without the intervention of users. Thirdly, after preprocessing, classification was undertaken. Image data dimension was reduced base on the principal component analysis (PCA). PCA reduces the data set to vectors that contain all the information of the data set. The vectors are used as the basis for the classification. In most PCA analyses, the most useful information is contained in the first six vectors. In this analysis, two vectors or two principals explained more than 95% of data variance.

An algorithm is then derived from the PCA vectors. The transformed image data with the reduced dimension was clustered by the K-means clustering method. The clustered images were segmented by the neighborhood information. This region-based segmentation is implemented by labeling connected components. The results of the K-means clustering were visualized graphically [20–22].

Finally, a nonparametric K-nearest neighbor (KNN) classifier was developed and applied for diagnosing each resulting segment [20–22]. The results of the KNN were visualized graphically. Areas the KNN algorithm diagnosed as high- and low-grade lesions were depicted in red.

**Results**

**Patient characteristics**

Forty-six patients were measured; 33 in Houston and 13 in Vancouver. The age of the patients ranged from 20 to 70 years, with a mean of 36 years and a median of 34 years old. The ethnic distribution showed that 3% were Asian, 18% were black, 18% were Hispanic, and 62% were white. One patient was excluded from the study because she did not have an MDC measurement or a LEEP treatment. Among participants, 12% were ex-smokers, 39% were current smokers, and the remaining 49% were never smokers. The menopausal status was as follows: 18% were post-menopausal, no patients were peri-menopausal, and 82% were pre-menopausal. Of the post-menopausal patients, none were currently taking hormones, one patient had stopped HRT 1 year ago. Of the pre-menopausal patients, 15% were using the oral contraceptive pill, 4% were using the contraceptive patch, 7% were using Depo-provera, and the remaining patients used condoms or had undergone a bilateral tubal ligation.
Image characteristics

The first 10 patients, at each institution, had measurements that did not image well. The camera was changed to a high-sensitivity camera that allowed detection of lower signal. This proved to be necessary for the imaging of CIN. Here, we report findings on 23 patients in Houston and 13 in Vancouver measured with the more sensitive camera. Each patient was measured at baseline (pre-acetic acid) with white light, 345 nm excitation, and 440 nm excitation from the xenon arc lamp, and after the placement of acetic acid these three measurements were repeated.

Figs. 5, 6, and 7 show the MDC images of 3 patients in Houston and Fig. 8 in Vancouver. For patients in Houston, we also include the histopathologic map and the K-means clustering algorithmic map of the cervix. In Figs. 5–7, one sees three MDC images. On the left are baseline (pre-acetic acid) images and on the right are the post-acetic-acid measurements. The first images are white light, the second 345 nm excitation pre- and post-acetic acid, and the third 440 nm excitation pre- and post-acetic acid. The lower left image is the K-means clustering results and the middle the KNN result selecting the high-grade lesion. The histopathologic map is on the lower right side. The Vancouver patient in Fig. 8 is imaged pre- and post-acetic-acid placement. Additionally, the figure is adjusted to show the effects of post-processing increases in contrast.

Fig. 5 shows a patient with a CIN2/HGSIL lesion over the os. The lesion is visualized with white light imaging in the post-acetic-acid image. At 345 nm excitation, that area appears darker blue. At 440 nm excitation, that area appears lighter green. The histopathologic map shows the CIN2/
HGSIL in the area imaged around the os. The KNN algorithm also shows the area; the image is larger and thus one must see the os in the K-means clustering to visualize the area. The KNN algorithm correctly identified the area of high-grade SIL.

Fig. 6 shows a patient with CIN2–3/HGSIL. The post-acetic-acid white image shows a diffuse lesion without much vascular atypia. The 345-nm excitation does not show a darker blue image like Fig. 5; however, the 440-nm excitation shows the colposcopically abnormal area as lighter green as in the previous example. The histopathologic map shows areas in red that represent CIN2–3/HGSIL; however, the blue areas show the areas where epithelium was denuded or separated from the stroma in processing. The K-means clustering image helps identify the landmark of the os. Again, the image is larger than the histopathologic map or the white light image. However, the KNN image fairly accurately matches the histopathologic map.

Fig. 7 shows 2 areas of LGSIL with prominent koilocytosis. The patient’s preprocedural Papanicolaou and cervical biopsy showed CIN2–3/HGSIL. In both the 345-nm excitation image and the 440-nm excitation image, the areas of LGSIL are red. We have noted this phenomenon in other images and believe it is probably due to porphyrin fluorescence. The K-means clustering algorithm identifies the os. The KNN algorithm accurately shows the lesion on the lower lip of the cervix, but the lesion anterior lip is not well visualized.

Fig. 8 shows one area of HGSIL at 5 o’clock in the white light colposcopy and in the histology (histology not shown). The image at 345 nm shows the area in darker blue before
the contrast is heightened as in Fig. 5. The image at 440 nm does not show as much contrast as desired between normal and abnormal areas as shown in Fig. 5.

Discussion

These findings are preliminary but promising. They resemble images obtained from a study of the hamster cheek pouch model of DMBA-induced carcinogenesis [19]. This work is at the earliest stages of device development. Fourteen years ago, point spectroscopy was at a similar level of development. Other investigators have found that fluorescence is a good tool for investigation of cervical dysplasia [13–15]. Our studies with the point probe agree and large diagnostic and screening trials of 850 and 1250 patients are nearly complete.

The current study shows that a more sensitive camera can improve the detection of low signal and thereby allow for images that correlate well with histopathologic mapping to be obtained. The current study also shows that histopathologic mapping can be hampered by missing epithelium and the lack of histopathologic information between sections. Other techniques for processing the histopathology of these ectocervical sections are being explored.

The K-means clustering algorithm produced images in which the os could be identified. Moreover, the KNN classifier allowed for identification of areas of both low-grade and high-grade SIL. K clustering and KKN classifiers are but one of several methods of analysis of image data.
We have also tried using the means, standard deviations, skewness factor, and kurtosis factor from measurements of the intensity and density of the red, green, and blue channels. We have also used that data transformed into the [yellow luminescence, chromatic blue, chromatic red] YCbCr color space. When the means, standard deviations, skewness factor, and kurtosis factor from measurements of the intensity and density of the YcbCr color space are used, promising algorithms can also be obtained [19].

The advantage of a device that images the whole field of the cervix and not the small 2-mm areas, like point spectroscopy, is that this device could be used for screening. While larger validation of the device and potential algorithm need to take place, we already envision a portable version of this instrument that can be battery powered. Theoretically, images could also be tele-transported and viewed by more experienced examiners, as is being done with colposcopic images [23,24].

In a future study we will conduct larger screening and diagnostic study to validate this device. We expect both improvements in the device, in the mapping, and in the algorithm. Lessons learned from the point probe studies will hopefully accelerate MDC development. Again, the goal is low-cost screening. The largest cost currently is that of the colposcope. The colposcope is of high quality but could be replaced by glasses. The portable system is under development and may cost only $2000–3000 per device. The use of screening fluorescence spectroscopy with low-cost devices in developing countries is imminent. The developing world is where the potential for decreasing mortality from cervical cancer is the highest.

Fig. 8. This figure shows pre- and post-acetic-acid images of the cervix. The figure shows a white light image like that of colposcopy, an image at 345 nm, and an image at 440 nm.
Acknowledgments

Support from the National Cancer Institute, Program Project Grant PO1 CA 82710-04 is gratefully acknowledged. We would also like to thank the providers: Judy Sandella, Alma Sbach, Karla Ballasteros, and Karen Rabel for their participation as providers.

References


