CHARACTERIZATION OF RAMAN SPECTROSCOPY FOR THE HUMAN CERVIX

By

Elizabeth Marie Kanter

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Engineering

August, 2008

Nashville, Tennessee

Approved:

Anita Mahadevan-Jansen

Duco E. Jansen

Gautam Rao

Yu Shyr

Rick Haselton

ACKNOWLEDGEMENTS

Over the past three years, my life has changed in a variety of ways; I have grown as a person both intellectually and emotionally. I will always have the fondest memories of my time at Vanderbilt University. There are several people I would like to thank for helping making this educational journey a wonderful one.

Anita Mahadevan-Jansen- Thank-you for being a wonderful advisor. You have challenged me and pushed me to work at the best of my ability. You have had a huge impact on my life and I really appreciate everything you have done for me.

Members of my committee- I would thank all the members of my committee. They have been very supportive and have all offered wonderful advice.

Members of the Biomedical Optics Lab- Thank you for being great colleagues and friends.

Shovan Majumder – Thank you for being there for me every step of the way. I really appreciate all the data processing you have done and of course the friendly conversations and all the wisdom you have passed on.

Matt Keller- Thank you for being the best friend anyone could have ever asked for. Thank you so much for giving me someone to bounce ideas off of and someone to talk to when I needed a friend.

Chetan Patil- Thanks so much always being willing to lend a helping hand. I don't know what I would have done without you.

Elizabeth Vargis- Thanks for taking over this project. I know there is so much potential in this project and I am so happy that some else who is excited about this project as I am is taking it over. Thanks so much for all your help over the past few months.

TriState Women's Health- I would like to thank everyone at TriSate Women's health. The doctors and nurses made my life so much easier by showing interest in my research and making sure that I could collect some patient data.

Vanderbilt Medical Center- I would like to thank all the doctors and nurses that put up with me over the years.

My Family and Friends- Thanks for all the support and love you have given me during my time in graduate school, without you I have no idea where I would be now. I would especially like to thank my mom, dad, my brother Eric and my boyfriend Steve Bartz.

Dad- Thanks for all your support. Throughout this entire procedure my goal was to be just like you… to be Dr. Kanter. You have always been my role model and the reason I pursued Biomedical Engineering. Without you I am not sure I would have ever made it through this process. I love you Daddy!

Funding Sources- Finally, I would like to thank my funding source, NIH R01-CA95405

TABLE OF CONTENTS

LIST OF TABLES

Table

Page

LIST OF FIGURES

Figure

Page

CHAPTER I

INTRODUCTION

1.1. Motivation

Cervical cancer is the second most common malignancy among women worldwide. It is estimated that in 2008, 3,870 deaths will occur in the United States alone from this disease and 11,070 new cases of invasive cervical cancer will be diagnosed. [1] It is believed that the incidence of *pre-invasive* squamous carcinoma of the cervix has risen dramatically in recent years, especially among women under the age of 50. [2] When cervical cancers are detected early, they are highly curable. In fact, early detection of cervical pre-cancer has played a central role in reducing the mortality associated with this disease in the United States over the last 50 years. [3] However, this trend is completely negated in countries such as Zambia where the mortality and prevalence rate of invasive cervical cancer is the second highest in the world. [4] Existing screening and detection techniques, the Pap smear and colposcopy, have several deficiencies that prevent efficient management of an otherwise controllable disease. An automated diagnostic method with improved sensitivity and specificity that could allow for a “See and Treat” protocol would significantly improve the management of the disease. The impact of such an approach would be particularly useful in situations where professional care is difficult to achieve.

1.2. Specific Aims

Specific Aim (1): Determine if the Raman spectra from low grade cervical lesions can be correctly differentiated from normal, inflammation, metaplasia and high grade lesions.

In previous studies, Raman spectroscopy was used to predict the presence of high-grade lesions and was evaluated using histology as the gold standard and compared with the sensitivity of colposcopy. In this aim, the focus was two-fold. (1) Raman spectra were acquired from patients with low-grade lesions in order to characterize their spectral signatures. Pre and peri menopausal patients were selectively recruited. (2) Using the spectra acquired in this study and previously acquired data; diagnostic algorithms based on non-linear regression methods were evaluated to determine their effectiveness. Menopausal status and location in the menstrual cycle were incorporated in the algorithms (aim 2).

Specific Aim (2): Characterize the Raman spectra of the normal cervix.

The second clinical study was designed to measure Raman spectra from areas of normal cervix in patients without evidence of cervical dysplasia. In this aim, multiple spectra of normal appearing areas were measured in each patient prior to hysterectomy and as well as in patients undergoing routine Pap smear. Each measurement was confirmed as normal by either the results of the biopsy from the excised hysterectomy tissues or by a normal pap smear for the Pap smear patients. Inter and intra-patient variability, the effects of menopausal state and location in the menstrual cycle on the Raman spectra acquired was assessed.

Specific Aim (3): Study the basis of observed spectral differences using confocal Raman Spectroscopy.

In developing Raman spectroscopy as a diagnostic tool for cervical precancers, the scientific basis for the success of the technique needs to be understood. What are the biochemical constituents of cervical tissues that are responsible for the observed Raman features? What causes the spectral differences between normal, non-precancerous and precancerous tissues that allow their differentiation? Confocal Raman spectroscopy will be used to develop this understanding on a microscopic as well as macroscopic level.

1.3 Summary of Chapters

This thesis consists 5 chapters after this introduction chapter; Chapter 2 is an extensive background on both cervical cancer and Raman Spectroscopy. The third chapter discusses the development of a new multi-class algorithm that shows better classification than the previous used algorithms. The fourth chapter shows that using menopausal status classification can be improved compared to what is seen in chapter 3. The fifth chapter discusses a large clinical study that shows the application of Raman spectroscopy in two different settings, including “see and treat” and colposcopy guided by optical biopsy. The last chapter (6) discusses the major conclusions found in this body of work and my contributions to the field. This work resulted in several smaller contributions, which are shown in the appendices. In the first appendix, the variation in the instrumentation used is discussed and how this affects the use of Raman spectroscopy as a diagnostic tool. This second appendix, discusses Aim 3. Due to the small number

of tissue samples that were collected due to the lack of availability, this was only as included as a part of the appendix.

1.4 References

1. American Cancer Society, "Cervical Cancer Resource Center", 2008.

2. L. Burke, D. A. Antonioli and B. S. Ducatman, *Colposcopy, text and atlas*, Appleton and Large, Norwalk, 1991.

3. G. H. Anderson, D. A. Boyes, J. L. Benedet, J. C. Leriche, J. P. Matisic, K. C. Suen, A. J. Worth, A. Millner and O. M. Bennett, "Organization and Results of the Cervical Cytology Screening-Program in British-Columbia, 1955-85," *British Medical Journal*, 296(6627), 975-978, 1988.

4. G. P. Parham, V. V. Sahasrabuddhe, M. H. Mwanahamuntu, B. E. Shepherd, M. L. Hicks, E. M. Stringer and S. H. Vermund, "Prevalence and predictors of squamous intraepithelial lesions of the cervix in HIV-infected women in Lusaka, Zambia," *Gynecologic Oncology*, 103(3), 1017-1022, 2006.

CHAPTER II

BACKGROUND

2.1. Normal Cervical Biology

The cervix measures 3-4 cm in length and is approximately 2.5cm in diameter. The size and shape of the cervix varies depending on age, parity and hormonal status [1], a typical normal cervix is shown in figure 2.2.a. The cervix is covered by two types of epithelia: the multi-layered squamous epithelium covers most of the ectocervix and is separated from the stroma by the basal layer. The columnar epithelium consists of a single layer of columnar cells, and covers the surface of the endocervical canal. The ectocervix consists of 15-20 layers of cells, has a large amount of glycogen, very few nerve endings and is typically pale pink in color. The endocervix on the other hand, is a single cell layer thick, has extensive sensory nerve endings, very little glycogen content and usually appears reddish in color. [1]

The interface of the two epithelia is called the squamous-columnar junction. Over time, the columnar epithelium is replaced by squamous epithelium, which causes the squamous-columnar junction to move towards the opening of the cervix called the os. This transitional epithelium is called *squamous metaplasia.* [2] Squamous metaplasia is an irreversible process shown in figure 2.1. The transformation zone (TZ) is the region where the columnar epithelium has been replaced or is currently being replaced by the new squamous epithelium. In younger women, the transformation zone is visible but as a woman ages the cervix shrinks due to decreased estrogen levels, the transformation

zone may only be partially or not at all visible. This region (transformation zone) is where most new cervical dysplasia occurs. [1]

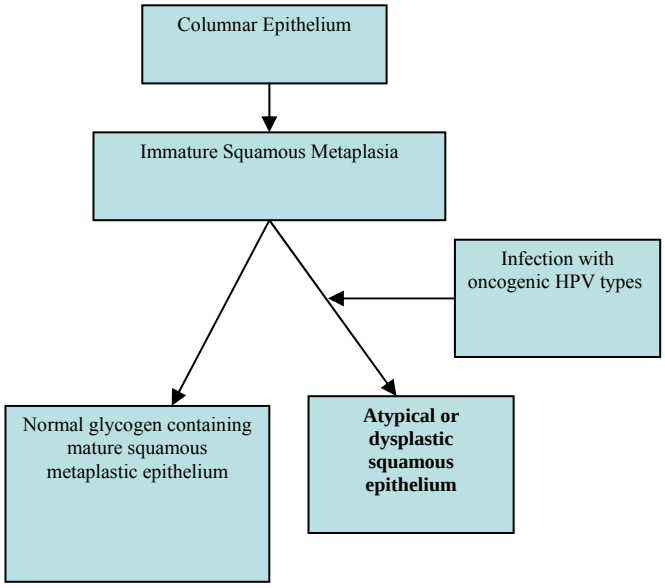


Figure 2.1. Progression of Columnar epithelium to squamous epithelium. [1]

2.2. Cervical Disease and Progression

Cervical intraepithelial neoplasia (CIN) refers to the development of neoplasia arising from the epithelium of the cervix. CIN refers to the precancerous stages of cervical carcinoma and is often also referred to as cervical dysplasia or squamous intraepithelial lesion (SIL). The progression and classifications can be seen in figure 2.2. Precancers may be categorized as mild, moderate and severe precancers (or dysplasia). The next step in the progression of this disease is carcinoma-in-situ (CIS) which is one step before the transformation of the dysplasia (precancer) to cancer [3-5]. Clinically speaking, cervical lesions can be divided into low grade lesions (mild precancers) and high grade lesions (moderate and severe precancers and CIS). It has been observed that

some cases of dysplasia regress and return to normal while other cases persist and develop into CIS and potentially cancer. [1] The distinction between high and low grade is important because patients with low grade lesions are typically followed but not treated whereas patients with high grade lesions are usually treated immediately with extended follow-up.

Benign	***Low grade (LGSIL)***	***High grade (HGSIL)***
Normal	HPV	Moderate CIN
Inflammation	Mild CIN	Severe CIN
Metaplasia		CIS

Figure 2.2: Progression of Cervical Disease (a) normal/benign cervix (Courtesy of the University of Washington) (b) mild dysplasia with arrow pointed to abnormal area that turned white due to application of acetic acid (courtesy of the Military Obstetrics & Gynecology) and (c) high grade cervical disease.

Human papilloma viruses (HPV) are viruses that predominantly infect skin and mucosal membranes and produce characteristic epithelial proliferation, which may undergo malignant transformations. Most HPV infections have no symptoms and go away over the course of a few years. [4] Similarities observed in the morphological changes of the epithelial cells between those induced by HPV and precancers has led to the suggestion that certain strains of HPV may be involved in the early stages of cervical

precancer and other strains may aid in the progression of the disease. [6] Thus HPV is typically placed in the same category as mild precancers (low grade lesions) and are clinically treated as such. A HPV vaccine (released in fall 2006), Gardasil™, has been shown to prevent infection from two types of HPV - HPV 16 and 18 - that together are known to cause 70 percent of cervical cancer cases worldwide. Gardasil™ also protects against HPV 6 and 11, which account for 90 percent of cases of genital warts. This vaccine is currently available to girls from ages 9-26 and is a sequence of 3 shots over a 6 month time period. However, females are not protected if they have been infected with any of the previously mentioned HPV type(s) prior to vaccination. Additionally, Gardasil™ does not protect against less common 12+ HPV types. Since the management and implementation of the vaccine is currently voluntary, there continues to be a need for the detection of precancerous changes in the cervix. [7]

Cervical dysplasia can occur anytime after a female becomes sexually active since most but not all cases of cervical dysplasia are brought on by sexually transmitted HPV infection. There are no symptoms of cervical dysplasia so screening is necessary for disease prevention. Cervical cancer on the other hand does not usually occur in women before the age of 40. Women 65 to 70 years of age who have had at least three normal Pap tests and no abnormal Pap tests in the last 10 years depending on the advice of their doctor may stop having annual Pap smears. [7] Therefore most cervical dysplasia occurs before women go through menopause unless there has been a prior abnormal Pap smear but most cervical cancer develops in women who are either peri-menopausal or menopausal with some exceptions.

2.3. Menopause and the Menstrual Cycle

Menopause is defined as the permanent cessation of the menstrual cycle of the female reproductive system. Perimenopause is defined as the transitional period from normal menstrual periods to no periods at all. During menopause, the ovarian source of estrogen declines (figure 2.3). This results in several physiological changes, such as thinning of the vaginal epithelium, decreased vaginal secretions and vascular instability. [8] The transition can, and usually does, take up to ten years and is associated with hormonal, physical and psychological changes in the person which affects the biochemical makeup of the cervix and therefore the spectral signature of cervical tissue.

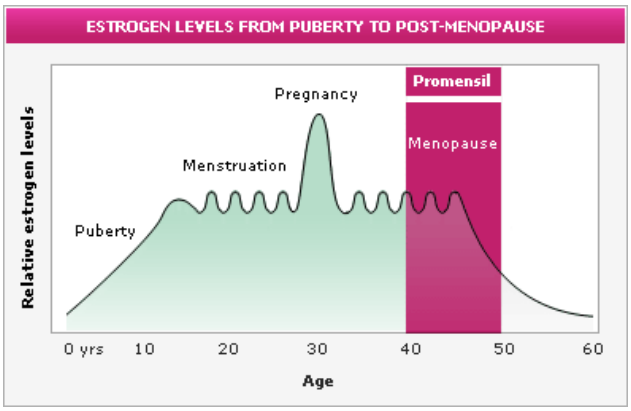


Figure 2.3. Estrogen levels from puberty to menopause (© 2006 Odiidis)

During menopause the estrogen levels decline rapidly but from puberty to menopause these levels fluctuate on a 28 day cycle (figure 2.4.). Estrogen is responsible for the maintenance and maturation of the uterus, fallopian tubes, cervix and vagina. During each cycle, estrogen regulates the consistency and composition of the cervical mucus and the cervix goes through many subtle changes. For example, as estrogen levels rise they trigger gradual opening of the os (leading to a softer cervix) and the begin production of

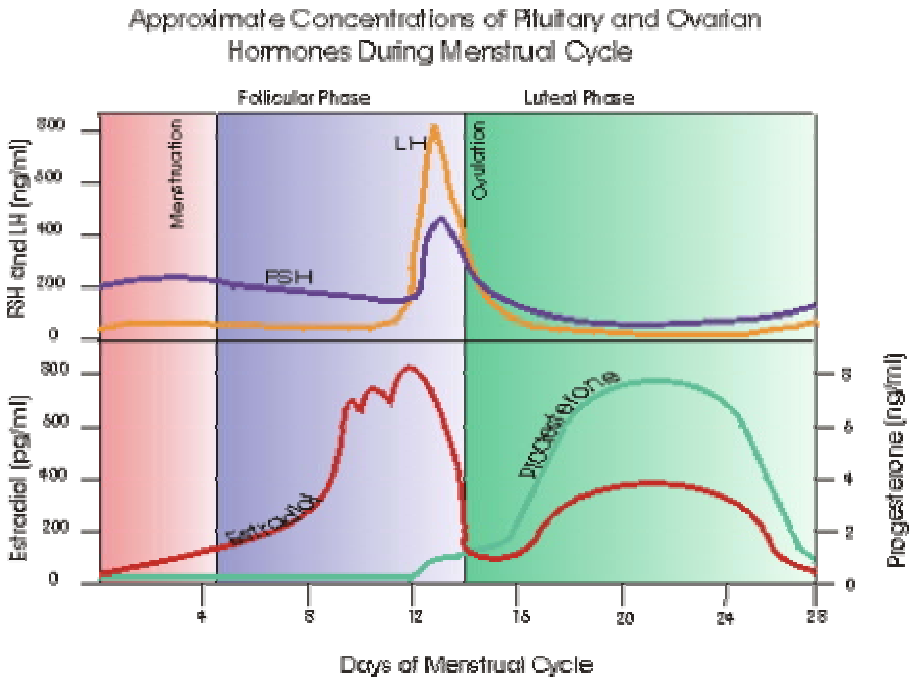


Figure 2.4. Hormone levels during the menstrual cycle. (© 2000 Molson Medical Informatics)

watery and elastic mucus. [8] There are 4 different phases in the menstrual cycle; the proliferative phase (before ovulation), ovulation, the secretory phase (after ovulation) and menstruation. Estrogen is present in increasing quantities during the proliferative phase; this phase is approximately 11 days long. Estrogen levels fluctuate after ovulation due to the change in the source of the estrogen. The estrogen in the proliferative phase comes from the developing follicles, while the estrogen in the secretory phase comes from the corpus luteum. The secretory phase is approximately 12 days long. Menstruation occurs when the corpus luteum stops making sex hormones and the levels drop, it typically lasts between 4 and 7 days. [8]

2.4. Current Screening Methods for Cervical disease

The primary screening tool for cervical precancer is the Papanicolaou (Pap) smear, where scrapings from the walls of the ecto- as well as endocervix, which contain a variable number of cells are examined and diagnosed. [9] Although the widespread application of the Pap smear as a screening tool has greatly decreased the incidence of cervical cancer [10], sampling and reading errors lead to high false positive and negative rates. A meta-analysis of the accuracy of Pap smears showed that in low prevalence populations, the mean sensitivity and specificity of the Pap smear was 48% and 95%, respectively. [11] Yet, another meta-analysis which did not restrict their analysis to low prevalence populations found a mean sensitivity (ability to correctly classify disease) and specificity (ability to correctly classify benign cervix) of 58% and 69%, respectively. [12] This suggests that the presence and extent of precancer is often wrongly estimated. Annual Pap smear is the standard of care amongst women in the developed countries, but for most women in developing populations like Asia and Africa Pap smear is not standard care and the disease is often not found until it is in its later stages. Thus a tool that can "See and Treat" would relieve a tremendous burden in the developing countries.

Colposcopy, which usually follows an abnormal Pap smear, is then typically used to direct the taking of biopsies. [13] A colposcope usually consists of a fixed focal length microscope with a variable magnification (4-40X) that is used to observe the surface of the cervix. The colposcopic image is produced by illuminating both the surface epithelium and the underlying stroma. To enhance the colposcopic image, 4-6% acetic acid is sometimes applied to the cervix. The application of acetic acid causes abnormal sites on the cervix to turn white, the area on figure 2.2.b is pointing to an area that turned

white after the application of acetic acid. This is a reversible process in that as the acetic acid depletes, the change is reversed to normal. This technique can be used repetitively on the same patient without harm. Colposcopy has a comparatively high degree of accuracy in pin-pointing the area and grade of the lesion. [14] However this procedure requires extensive training and its accuracy is variable and limited even in the hands of expert practitioners. [15] As a result, colposcopy is not used as a stand-alone method of diagnosis. Once the abnormal sites are identified colposcopically, multiple biopsies are taken using a standard punch biopsy forceps. These tissue samples are fixed in formalin and then sent for histologic examination. Histology then forms the gold standard for diagnosis and determination of treatment.

Several new techniques for cervical disease detection have been introduced in the past few years, such as wet prep and cervicography. Wet prep, or *liquid-based thin-layer slide preparation, modifies* conventional Pap smear. This technique rinses the cells into a vial of liquid instead of smearing them onto a slide, making the final slide easier to read by reducing the clumping of cells. Results of this new method suggest that it is comparable to, or more sensitive than (possibly reducing reading error), the conventional Pap smear for the detection of abnormalities. [7] Wet prep is now routinely used to aid the Pap smear at least in the United States. In cervicography, a high resolution photograph is taken of the cervix after the application of acetic acid and sent to a central laboratory to be read by colposcopists who have received specialized training in interpretation of these photographs. This technique, although still in the research phases, could be used in junction with or replace Pap smear. Cervicography would eliminate the need for a trained colposcopists to look at each individual cervix and would standardize

how these images are read [16]. The limitation to this method is the extended amount of time needed to send the photographs off to be read.

Thus existing conventional screening and diagnostic techniques for cervical precancers have several deficiencies that prevent efficient management of an otherwise controllable disease. Standard of care ultimately continues to rely on histology to make a definitive diagnosis before treatment is planned. An accurate, automated diagnostic method could allow faster, more effective patient management.

2.5. Optical Methods

The application of optical methods is suggested because it can detect alterations in tissue architecture and biochemical composition associated with the progression of disease. [17-28] Optical methods can provide automated, fast and non-invasive characterization of normal and non-normal tissues *in vivo*. Several different optical methods have been used to characterize the cervix due to the need to monitor disease development and progression.

2.5.1 Optical Coherence Tomography (OCT)

OCT is a high resolution, cross sectional imaging modality analogous to ultrasound. It uses a low coherence near infrared light source to obtain depth-resolved images of tissue microstructure. Even in highly scattering tissues, structures up to 2mm deep can be imaged. [24] OCT has high spatial resolution (~11um dependent on source) and images can be taken in real-time. A recent study done by the Cleveland Clinic recruited 220 patients that had a history of an abnormal Pap smear (both in the US and

the Dominican Republic). On each patient, they completed a visual exam, a colposcopic exam and OCT. Using visual inspection, colposcopy and OCT together on abnormal tissue has a sensitivity of 46% and a specificity of 69%. [20] In figure 2.5, OCT images from this study of the cervix are displayed. In figure 2.5.a an OCT image of a cervix with high grade dysplasia is shown, and in figure 2.5.b an OCT image of a normal cervix is shown. To the naked eye there seems to be little to no difference between these two OCT images.

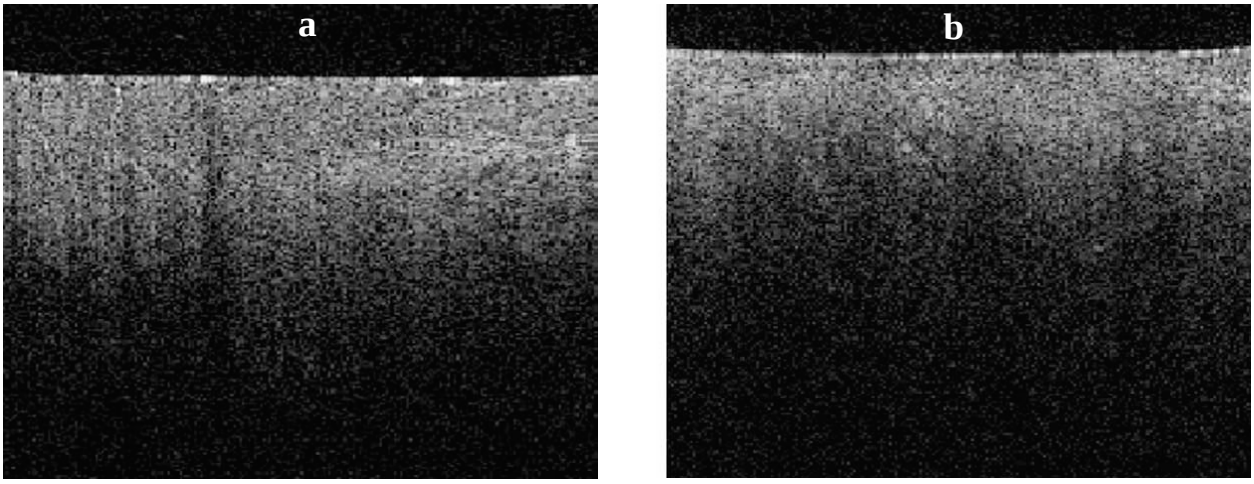


Figure 2.5: OCT images of the (a) CIN III and (b) unstructured normal ecto-cervix. [20]

Since, a technique such as OCT is ultimately dependent on structural changes in the tissue microstructure and since other indications are known to have similar architectural disruptions, the reported study suggests OCT does not provide the performance needed to improve cervical precancer detection.

2.5.2 Fluorescence Spectroscopy

Fluorescence spectroscopy is a commonly tested optical technique for the *in vivo* detection of diseases in general and cancers in particular. Fluorescence spectroscopy of

both exogenous and endogenous chromophores has been successfully used to identify neoplastic cells and tissues in a variety of organ systems. [27] Fluorescence spectroscopy has been studied extensively for the diagnosis (and potential for screening) of cervical precancers. [17, 18, 29, 30] The first *in vitro* studies to assess the potential of fluorescence spectroscopy for cervical precancers was done by Mahadevan et. al [25]. This *in vitro* study highlighted the inter-patient variability of fluorescence signals and suggested the need for paired analysis (comparison of abnormal to the normal of the same patient).

Richards-Kortum *et al* have since extensively developed and evaluated fluorescence spectroscopy for the detection of cervical precancers *in vivo.* [31] Early efforts were primarily focused on the application of single point fluorescence spectra acquired at multiple excitation wavelengths for the detection of cervical lesions. [32] Multivariate discriminations algorithms were developed based on fluorescence spectra acquired from 95 patients at three excitation wavelengths (337, 380 and 460 nm) and tested. The prospective sensitivity and specificity of the multivariate algorithms based on paired fluorescence information for differentiating cervical precancers from normal tissues was 82% and 73%. The same studies described above show that the specificity in discriminating precancers from *non-precancerous* tissues is reduced to 68%. [33] Subsequent studies by this group have since moved to the measurement of entire excitation-emission matrices in an attempt to improve the diagnostic performance. More recently the use of diffuse reflectance has been additionally included in a continued effort to enhance the performance of fluorescence spectroscopy for cervical precancer

detection. They have shown that using a multispectral digital colposcope to collect diffuse reflectance, images they can achieve 79% sensitivity and 88% specificity. [34]

Another group from the University of Alabama- Birmingham found that they could detection high grade lesions compared to all other tissue with a sensitivity of approximately 90% and a sensitivity of 50%. [23] This result along with similar observations made in other organs indicate that fluorescence spectra of precancerous and cancerous tissues of, for example, the cervix and colon are similar in many patients to the fluorescence spectra of benign abnormalities such as inflammation, hyperplasia and metaplasia. [33, 35] This suggests that the method of fluorescence diagnosis yields an unacceptably high false positive rate in discriminating cancers and precancers from all other tissues. It should also be noted that clinically it is extremely important to detect every high grade lesion. Thus, a sensitivity of 79% in detecting high grade lesions while presenting a significant improvement in current practice is not sufficient to change the standard of care which relies on biopsies and histology for treatment. [33]

Several studies have been done to improve the sensitivity and specificity of the fluorescence measurements by taking menopausal state and location in the menstrual cycle into consideration. A study done by Gill *et. al*., showed that there was a statistical difference between pre and post menopausal fluorescence data and that the post menopausal women have a higher average fluorescence signal and this difference suggests collagen cross-linking with menopause. [36] Another study done by Cox *et. al.*, looked at the effects of the menstrual cycle on fluorescence measurements. The study indicates that main concern of the menstrual cycle is hemoglobin absorption and can be avoided if measurements are not taken in the first eight days of the cycle. [37] These

reports clearly indicate the need to include the menopausal status as a variable in the development of diagnostic algorithms.

2.6 Raman Spectroscopy

Raman spectroscopy has been used for many years to probe into the biochemistry of various biological molecules. [26] In recent years, there has been interest in using this technique in diagnostics. [22] Raman spectroscopy probes different characteristics of materials than fluorescence. Raman scattering is an inelastic scattering process, which arises from perturbations of the molecule that induces vibrational or rotational transitions. [38] Thus Raman spectroscopy is a molecular specific technique that can be used as a biochemical tool for study of different materials; in particular this technique has the capability to provide differential diagnosis of precancers and cancers.

Although, only a limited number of biological molecules contribute to tissue fluorescence, most with broadband emission, several types of biological molecules such as nucleic acids, proteins and lipids have distinctive Raman features that yield molecular specific structural and environmental information. These molecules have been studied in solutions as well as in their natural microscopic environment. [28] Results indicate that the molecular and cellular changes that occur with cancer result in distinct Raman spectra. The transitional changes in precancerous tissues as well as in benign abnormalities such as inflammation could also yield characteristic Raman features that allow their differentiation. For example, one of the more prominent changes that occur with cancer and precancer is increased cellular nucleic acid content; extensive DNA studies indicate that it may be possible to sample this change using Raman spectroscopy.

[21] On the basis of these biochemical differences, several groups have studied the potential of vibrational spectroscopy for cancer diagnosis in various organ sites. [26] These groups have shown that features of the vibrational spectrum can be related to molecular and structural changes associated with neoplastic transformation. Raman spectroscopy has been applied towards *in vitro* detection of cancers of epithelial and mesenchymal origin such as breast, colon, esophagus and gynecologic tissues. [22] Most recently, new reports have been published on the application of Raman spectroscopy for the detection of cancers *in vivo* in organs such as the cervix, skin breast and the gastrointestinal tract (GI). [19, 39-42]

2.6.1 Raman Theory

Raman spectroscopy is based on the Raman Effect, which occurs when energy is exchanged between the incident photon and the scattering molecules. Inelastic or Raman scattering occurs when a photon is incident on molecule and the scattered photon has more energy than the incident photon (Raman Anti-Stokes Scattering) or a photon is incident on molecule scattered photon has less energy than the incident photon (Raman Stokes Scattering). If the scattered photon and incident photon have the same amount of energy, it is referred to as Rayleigh or Elastic Scattering. A summary of these interactions can be seen in figure 2.6.

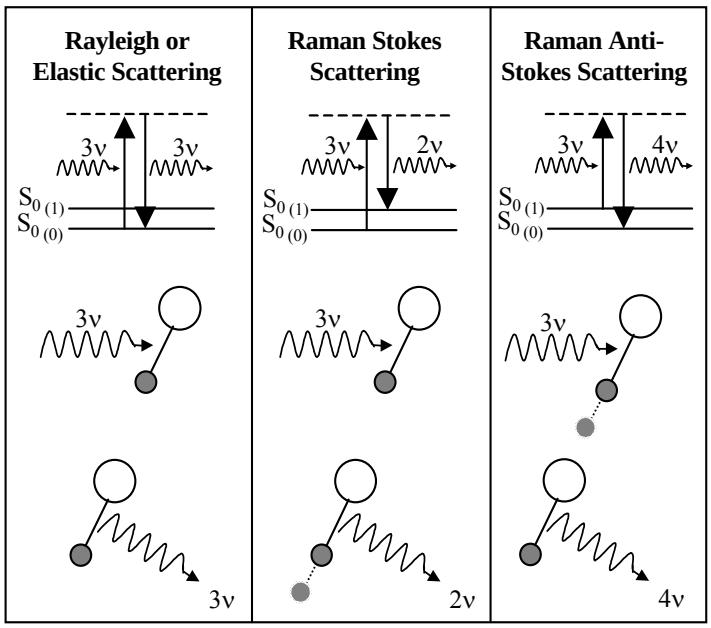


Figure 2.6. Schematic of the different types of scattering. [43]

Raman spectra are typically plotted as the collected scattered intensity vs. Raman shift (frequency shift) between the incident and scattered photon. In the Raman systems used in this project, the incident photon is either 785nm or 830nm. Depending on the characteristics of the incident molecule, the frequency shift varies. Many different chemical bonds have a narrow Raman peak that can be associated that particular bonds but since tissue is complex (and has very weak Raman signal), it is not trivial to determines the origin of each Raman peak. Therefore the Raman spectra of cervical tissue needs to be extensively studied to determine where there are differences between disease classes as well as what might be contributing to these differences in order to determine the feasibility of this as a diagnostic tool. [44]

2.6.2 Using Raman Spectroscopy for Cervical Cancer diagnosis

The largest clinical Raman study of the human cervix to date has been done in our lab at Vanderbilt University. Raman spectra were measured *in vivo* from 110 patients undergoing diagnostic or therapeutic procedures involving the removal of cervical tissue prior to tissue excision. The Raman spectra were acquired using a portable Raman spectroscopy system consisting of a 785 nm diode laser (Process Instruments, Inc., Salt Lake City, UT), custom fiber optic probe, imaging spectrograph (Kaiser Optical Systems, Inc., Ann Arbor, MI), and back-illuminated, deep-depletion, charge coupled device (CCD) camera (Princeton Instruments, Princeton, NJ), all controlled with a laptop computer. The spectra were corrected for instrumentation variations, then noise smoothed using a Savitzky-Golay filter and then the fluorescence was subtracted using the modified polyfit technique with a 5th degree polynomial. The resulting spectra were then correlated with the corresponding histopathologic diagnosis to characterize the differences between various diagnostic categories.

The resulting mean spectra are shown in figure 2.7 (for 79 of the patients). The most consistent peaks are labeled and found at 1006 (phenylalanine), 1058, 1086 (Lipids, nucleic acids backbone), 1244, 1270 (proteins), 1324 (adenine), 1450 (lipids and proteins), 1550, 1655 (lipids) cm^{-1}. [45] Although the peak shapes and locations show consistency across all pathology classifications, there are small but significant differences in peak intensities between the different pathology categories.

Based on these differences observed and using histology as the gold standard, logistic regression discrimination algorithms were developed to distinguish between normal ectocervix, squamous metaplasia, and high grade dysplasia using independent training

and validation sets of data. The classification model was constructed to automatically classify spectra into one of two categories (high grade dysplasia or benign cervix) using a two-tiered logistic regression model. An unbiased estimate of the accuracy of the model indicates that Raman spectroscopy was able to distinguish between high grade dysplasia and benign areas of the cervix (normal ectocervix and squamous metaplasia) with sensitivity of 89% and specificity 81% while colposcopy in expert hands was able to discriminate with a sensitivity of 87% and specificity of 72%. [46]

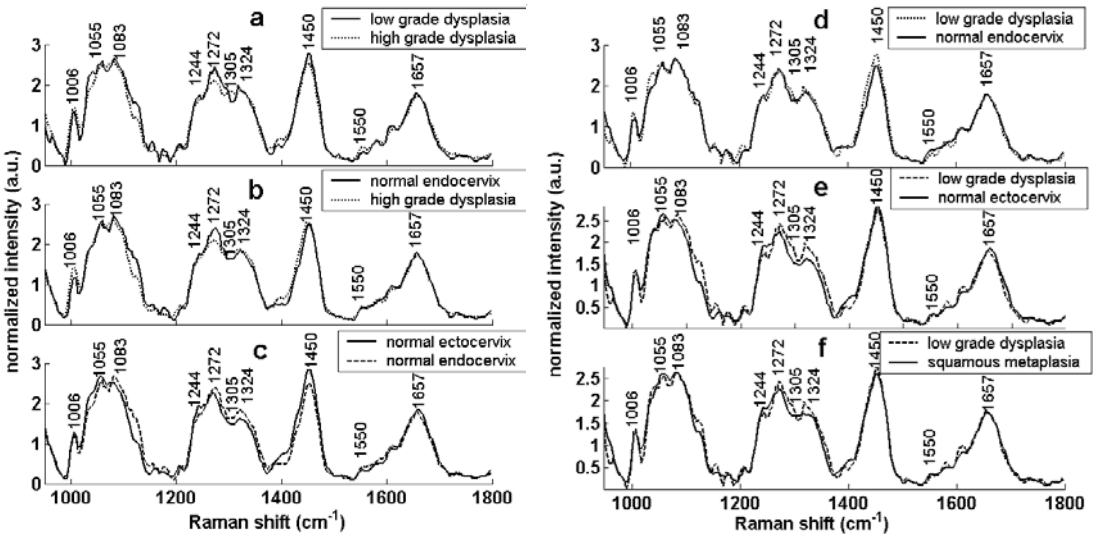


Figure 2.7**:** Mean Raman spectral overlays for the following categories: (a) high grade dysplasia (n=29 spectra) and low grade dysplasia (n=6 spectra), (b) normal endocervix (n=8 spectra) and high grade dysplasia, (c) normal ectocervix (n=100 spectra) and normal endocervix, (d) low grade dysplasia and normal endocervix, (e) low grade dysplasia and normal ectocervix, (f) low grade dysplasia and squamous metaplasia (29 spectra) [47].

The limitation of this particular study is that the discrimination algorithms developed above were binary and this did not allow for multiple classes that the tissue

could belong to. Thus, a new discrimination algorithm was developed based on novel statistical methods: Maximum representation and discrimination feature (MRDF) combined with sparse multinomial logistic regression (SMLR).

MRDF is a method of feature extraction whose objective is to maximally extract the diagnostic information otherwise hidden in a set of measured spectral data by reducing its dimensionality through a set of mathematical transforms. Given a set of input data comprising spectra from different classes with a given dimensionality, nonlinear MRDF aims to find a set of nonlinear transformations of the input data that optimally discriminate between the different classes in a reduced dimensionality space. It basically invokes nonlinear transforms (restricted order polynomial mappings of the input data) in two successive stages. A list of the terms used in this description on listed in table 2.1.

In the first stage, the input spectral data $\mathbf{x} = [x_1, x_2, ...x_N]^T$ (intensities corresponding frequency shift of the spectra) from each tissue type are raised to the power p' to produce the associated nonlinear input vectors $\mathbf{x}_{p'} = [x_1^{p'}, x_2^{p'}, ..., x_N^{p'}]$, which are then subject to a transform Φ'_M such that $\mathbf{y}'_M = \Phi'^T_M \mathbf{x}_{p'}$ are the first stage output features in the nonlinear feature space of reduced dimension $M << N$. In the second stage, the reduced M dimensional output features $\mathbf{y}'_M$ for each tissue type are further transformed nonlinearly to the power p to produce higher order features $\mathbf{y}'_{Mp} = [y'^p_1, y'^p_2, ..., y'^p_M]$, and a second transform Φ_K is computed so as to yield the final output features $\mathbf{y}_K = \Phi^T_K \mathbf{y}_{Mp}$ in the nonlinear feature space of dimension K ($K \leq M$). [48]

Table 2.1 MRDF terms

Term	Description
$\mathbf{x}$	Input vector – spectral intensities
N	Dimension of $\mathbf{x}$
p'	Normalization factor
$\mathbf{x}_{\mathbf{p}'}$	$\mathbf{x}$ is raised to the power p' to form a restricted polynomial
Φ'_M	Linear transform
$\mathbf{y'_m}$	Resultant vector when Φ'_M is applied to $\mathbf{x}_{\mathbf{p}'}$
M	Dimensionality of the vector after the first stage of MRDF
p	Normalization factor
$\mathbf{y'_{pm}}$	$\mathbf{y'_m}$ is raised to the power p to form a restricted polynomial
Φ_K	Linear transform
$\mathbf{y_k}$	Resultant vector when Φ_K is applied to $\mathbf{y'_{pm}}$
K	Dimensionality of the final vector in MRDF

SMLR is a method of supervised classification. It is a probabilistic multi-class model based on sparse Bayesian machine-learning framework of statistical pattern recognition. The central idea of SMLR is to separate a set of labeled input data into its constituent classes by predicting the posterior probabilities (likeliness) of their class-membership. It computes the posterior probabilities using a multinomial logistic

regression model and constructs a decision boundary that separates the data into its constituent classes based on the computed posterior probabilities following Bayes' rule i.e. a class is assigned to a data for which its posterior probability is the highest [48].

Now by combining Raman spectroscopy and the advanced statistical methods mentioned, the ability to successfully diagnosis cervical dysplasia will be shown. This body of work will discuss both the advantages and the disadvantages associated with this method for cervical dysplasia detection and make suggestions on what needs to be done in the future for this technique to be used regularly in the clinic.

2.7 References

1. R. S. John W. Sellors, *Colposcopy and Treatment of Cervical Neoplasia: A Beginners' Manual*, International Agency for Research on Cancer, Lyon, 2003.

2. K. E. Krantz, "The Anatomy of the Human Cervix, Gross and Microscopic," *The Biology of the Cervix*, The University of Chicago Press, Chicago, 1973.

3. J. L. H. Evers and M. J. Heineman, *Gynecology - A Clinical Atlas*, CV Mosby Co, St.Louis, 1990.

4. T. Wright, R. J. Kurman and A. Ferenczy, "Cervical Intraepithelial Neoplasia," *Blaustien's Pathology of the Female Genital Tract*, Springer-Verlag, New York, 1994.

5. I. Ramzy, *Essentials of Gynecologic and Obstetric Pathology.*, Appleton - Century - Crofts, Norwalk, 1983.

6. D. V. Coleman, C. Wickenden and A. D. B. Malcolm, "Association of Human Papillomavirus with Squamous Carcinoma of the Uterine Cervix," *Ciba Foundation Symposia*, 120(175-189, 1986.

7. American Cancer Society, "Cervical Cancer Resource Center", 2007.

8. L. S. Castanzo, *2nd Edition Physiology*, Saunders, Philadelphia, 2002.

9. B. Bates, *A Guide to Physical Examination*, J.B. Lippincott Co, Philadelphia, 1974.

10. E. R. Myers, D. C. McCrory, S. Subramanian, N. McCall, K. Nanda, S. Datta and D.

B. Matchar, "Setting the target for a better cervical screening test: Characteristics of a cost-effective test for cervical neoplasia screening," *Obstetrics and Gynecology*, 96(5), 645-652, 2000.

11. K. Nanda, D. C. McCrory, E. R. Myers, L. A. Bastian, V. Hasselblad, J. D. Hickey and D. B. Matchar, "Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: A systematic review," *Annals of Internal Medicine*, 132(10), 810-819, 2000.

12. M. T. Fahey, L. Irwig and P. Macaskill, "Metaanalysis of Pap Test Accuracy," *American Journal of Epidemiology*, 141(7), 680-689, 1995.

13. L. Burke, D. A. Antonioli and B. S. Ducatman, *Colposcopy, text and atlas*, Appleton and Large, Norwalk, 1991.

14. U. Minoru, *Cervical Adenocarcinoma: A Colscopic Atlas*, Ishiyaku - EuroAmerica Inc, St Louis, 1985.

15. M. F. Mitchell, "Accuracy of Colposcopy," *Consultations in Obstetrics and Gynecology, vol. 6*, 1994.

16. S. Bomfim-Hyppolito, E. S. Franco, R. G. D. Franco, C. M. de Albuquerque and G. C. Nunes, "Cervicography as an adjunctive test to visual inspection with acetic acid in cervical cancer detection screening," *International Journal of Gynecology & Obstetrics*, 92(1), 58-63, 2006.

17. S. K. Chang, I. Pavlova, N. M. Marin, M. Follen and R. Richards-Kortum, "Fluorescence spectroscopy as a diagnostic tool for detecting cervical pre-cancer," *Gynecologic Oncology*, 99(3), S61-S63, 2005.

18. S. K. Chang, Y. N. Mirabal, E. N. Atkinson, D. Cox, A. Malpica, M. Follen and R. Richards-Kortum, "Combined reflectance and fluorescence spectroscopy for in vivo detection of cervical pre-cancer," *Journal of Biomedical Optics*, 10(2), 2005.

19. M. V. P. Chowdary, K. K. Kumar, J. Kurien, S. Mathew and C. M. Krishna, "Discrimination of normal, benign, and malignant breast tissues by Raman spectroscopy," *Biopolymers*, 83(5), 556-569, 2006.

20. P. F. Escobar, L. Rojas-Espaillat, S. Tisci, C. Enerson, J. Brainard, J. Smith, N. J. Tresser, F. I. Feldchtein, L. B. Rojas and J. L. Belinson, "Optical coherence tomography as a diagnostic aid to visual inspection and colposcopy for preinvasive and invasive cancer of the uterine cervix," *International Journal of Gynecological Cancer*, 16(5), 1815-1822, 2006.

21. M. S. Feld, R. Manoharan, J. Salenius, J. Orenstein-Carndona, T. J. Romer, J. F. B. III, R. R. Dasari and Y. Wang, "Detection and Characterization of Human Tissue Lesions with Near Infrared Raman Spectroscopy," *Proc. of SPIE*, 2388(1995.

22. E. B. Hanlon, R. Manoharan, T. W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari and M. S. Feld, "Prospects for in vivo Raman spectroscopy," *Physics in Medicine and Biology*, 45(2), R1-R59, 2000.

23. W. K. Huh, R. M. Cestero, F. A. Garcia, M. A. Gold, R. S. Guido, K. McIntyre-Seltman, D. M. Harper, L. Burke, S. T. Sum, R. F. Flewelling and R. D. Alvarez, "Optical detection of high-grade cervical intraepithelial neoplasia in vivo: Results of a 604-patient study," *American Journal of Obstetrics and Gynecology*, 190(5), 1249-1257, 2004.

24. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito and et al., "Optical coherence tomography," *Science*, 254(5035), 1178-81, 1991.

25. A. Mahadevan, M. F. Mitchell, E. Silva, S. Thomsen and R. R. Richardskortum, "Study of the Fluorescence Properties of Normal and Neoplastic Human Cervical Tissue," *Lasers in Surgery and Medicine*, 13(6), 647-655, 1993.

26. A. Mahadevan-Jansen and R. Richards-Kortum, "Raman Spectroscopy for the detection of cancers and precancers," *J Biomed Opt*, 1(1), 31-70, 1996.

27. N. Ramanujam, "Fluorescence spectroscopy of neoplastic and non-neoplastic tissues," *Neoplasia*, 2(1-2), 89-117, 2000.

28. J. Twardowski and P. Anzenbacher, "Raman and IR Spectroscopy in Biology and Biochemistry," Ellis Horwood, New York, 1994.

29. U. Utzinger, E. V. Trujillo, E. N. Atkinson, M. F. Mitchell, S. B. Cantor and R. Richards-Kortum, "Performance estimation of diagnostic tests for cervical precancer based on fluorescence spectroscopy: Effects of tissue type, sample size, population, and signal-to-noise ratio," *Ieee Transactions on Biomedical Engineering*, 46(11), 1293-1303, 1999.

30. J. S. Lee, O. Shuhatovich, R. Price, B. Pikkula, M. Follen, N. McKinnon, C. MacAulay, B. Knight, R. Richards-Kortum and D. D. Cox, "Design and preliminary analysis of a study to assess intra-device and inter-device variability of fluorescence spectroscopy instruments for detecting cervical neoplasia," *Gynecologic Oncology*, 99(3), S98-S111, 2005.

31. M. F. Mitchell, S. B. Cantor, N. Ramanujam, G. Tortolero-Luna and R. Richards-Kortum, "Fluorescence spectroscopy for diagnosis of squamous intraepithelial lesions of the cervix," *Obstetrics and Gynecology*, 93(3), 462-470, 1999.

32. A. Milbourne, S. Y. Park, J. L. Benedet, D. Miller, T. Ehlen, H. Rhodes, A. Malpica, J. Matisic, D. Van Niekirk, E. N. Atkinson, N. Hadad, N. Mackinnon, C. Macaulay, R. Richards-Kortum and M. Follen, "Results of a pilot study of multispectral digital colposcopy for the in vivo detection of cervical intraepithelial neoplasia," *Gynecol Oncol*, 99(3 Suppl 1), S67-75, 2005.

33. N. Ramanujam, M. F. Mitchell, A. MahadevanJansen, S. L. Thomsen, G. Staerkel, A. Malpica, T. Wright, N. Atkinson and R. RichardsKortum, "Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths," *Photochemistry and Photobiology*, 64(4), 720-735, 1996.

34. S. Y. Park, M. Follen, A. Milbourne, H. Rhodes, A. Malpica, N. MacKinnon, C. MacAulay, M. K. Markey and R. Richards-Kortum, "Automated image analysis of digital colposcopy for the detection of cervical neoplasia," *Journal of Biomedical Optics*, 13(1), 014029-10, 2008.

35. K. T. Schomacker, J. K. Frisoli, C. C. Compton, T. J. Flotte, J. M. Richter, N. S. Nishioka and T. F. Deutsch, "Ultraviolet Laser-Induced Fluorescence of Colonic Tissue - Basic Biology and Diagnostic Potential," *Lasers in Surgery and Medicine*, 12(1), 63-78, 1992.

36. E. M. Gill, A. Malpica, R. E. Alford, A. R. Nath, M. Follen, R. R. Richards-Kortum and N. Ramanujam, "Relationship between collagen autofluorescence of the human cervix and menopausal status," *Photochemistry and Photobiology*, 77(6), 653-658, 2003.

37. D. D. Cox, S. K. Chang, M. Y. Dawood, G. Staerkel, U. Utzinger, R. R. Richards-Kortum and M. Follen, "Detecting the signal of the menstrual cycle in fluorescence spectroscopy of the cervix," *Applied Spectroscopy*, 57(1), 67-72, 2003.

38. R. J. Colthrup, *Infrared and Raman spectroscopy*, 1991.

39. T. R. Hata, T. A. Scholz, I. V. Ermakov, R. W. McClane, F. Khachik, W. Gellermann and L. K. Pershing, "Non-invasive Raman spectroscopic detection of carotenoids in human skin," *Journal of Investigative Dermatology*, 115(3), 441-448, 2000.

40. M. G. Shim, L. Song, N. E. Marcon and B. C. Wilson, "In vivo near-infrared Raman spectroscopy: Demonstration of feasibility during clinical gastrointestinal endoscopy," *Photochemistry and Photobiology*, 72(1), 146-150, 2000.

41. A. Mahadevan-Jansen, W. F. Mitchell, N. Ramanujam, U. Utzinger and R. Richards-Kortum, "Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue in vivo," *Photochemistry and Photobiology*, 68(3), 427-431, 1998.

42. A. P. Oliveira, R. A. Bitar, L. Silveira, R. A. Zangaro and A. A. Martin, "Near-infrared Raman spectroscopy for oral carcinoma diagnosis," *Photomedicine and Laser Surgery*, 24(3), 348-353, 2006.

43. S. K. Chang, Y. N. Mirabal, E. N. Atkinson, D. Cox, A. Malpica, M. Follen and R. Richards-Kortum, "Combined reflectance and fluorescence spectroscopy for in vivo detection of cervical pre-cancer," *J Biomed Opt*, 10(2), 024031, 2005.

44. T. Vo-Hinh, *Biomedical Photonics Handbook*, CRC Press, Boca Raton, 2003.

45. R. A. Bitar, H. D. S. Martinho, C. J. Tierra-Criollo, L. N. Z. Ramalho, M. M. Netto and A. A. Martin, "Biochemical analysis of human breast tissues using Fourier-transform Raman spectroscopy," *Journal of Biomedical Optics*, 11(5), -, 2006.

46. A. Viehoever Robichaux, E. Kanter, H. Shappell, D. Billheimer, H. Jones and A. Mahadevan-Jansen, "Characterization of Raman Spectra Measured In vivo for the Detection of Cervical Dysplasia," *Applied Spectroscopy*, in review- 07-04719(2007.

47. A. Robichaux-Viehoever, E. Kanter, H. Shappell, D. Billheimer, H. Jones III and A. Mahadevan-Jansen, "Characterization and Quantification of the Sources of Variability Present in Raman Spectra of the Cervix," *Photochemistry and Photobiology*, in review-PHP-2007-01-RA-0021 (2007).

48. S. K. Majumder, E. Kanter, A. R. Viehoever, H. Jones and A. Mahadevan-Jansen, "Near-infrared Raman spectroscopy for in-vivo diagnosis of cervical dysplasia: a probability-based multi-class diagnostic algorithm", *Advanced Biomedical and Clinical Diagnostic Systems V,* SPIE, 6430(64300Q-11), 2007.

CHAPTER III

MULTI-CLASS DISCRIMINATION OF CERVICAL PRECANCERS USING RAMAN SPECTROSCOPY

3.1 Abstract

Raman spectroscopy has the potential to differentiate among the various stages leading to high-grade cervical cancer, such as normal, squamous metaplasia, and low-grade cancer. For Raman spectroscopy to successfully differentiate among the stages, an applicable statistical method must be developed. Algorithms like linear discriminate analysis (LDA) are incapable of differentiating among three or more types of tissues. We developed a novel statistical method combining maximum representation and discrimination feature (MRDF) to extract diagnostic information with sparse multinomial logistic regression (SMLR) to classify spectra based on nonlinear features for multi-class analysis of Raman spectra. We found that high-grade spectra classified correctly 95% of the time; low-grade data was classified correctly 74% of the time, improving sensitivity from 92% to 98% and specificity from 81% to 96% suggesting that MRDF with SMLR is a more appropriate technique for categorizing Raman spectra. SMLR also outputs a posterior probability to evaluate the algorithm's accuracy. This combined method holds promise to diagnose subtle changes leading to cervical cancer.

3.2 Introduction

Raman spectroscopy has been used for many years to probe into the biochemistry of various biological molecules.[1] It is a molecular specific technique that can be used

as a biochemical tool to provide differential diagnosis of precancers and cancers. Several biological molecules such as nucleic acids, proteins and lipids have distinctive Raman features that yield molecularly specific structural and environmental information. Results indicate that molecular and cellular changes that occur in precancerous tissues as well as in benign abnormalities, such as inflammation, yield characteristic Raman features that allow their differentiation. For example, one of the more prominent changes that occur with cancer and precancer is increased cellular nucleic acid content; extensive DNA studies indicate that it may be possible to detect this change using Raman spectroscopy.[2] On the basis of these biochemical differences, several groups have studied the potential of vibrational spectroscopy for cancer diagnosis in various organ sites.[1] These groups have shown that features of the vibrational spectra can be related to molecular and structural changes associated with neoplastic transformation. Accordingly, Raman spectroscopy has been applied towards *in vitro* detection of cancers of epithelial and mesenchymal origin such as breast, colon, esophagus and gynecologic tissues.[3] While many challenges have prevented the widespread application of Raman spectroscopy for disease detection, recent developments in detector and source technologies have resulted in acquisition of Raman spectra from tissue in 1-3 seconds. Several fiber optics probes have also been developed that are capable of measuring Raman spectra *in vivo* making it possible to apply this technique in a clinical setting.[4] There have been an increased number of reports published, on applying Raman spectroscopy for detecting cancers *in vivo*, such as in the cervix, skin, breast and the gastrointestinal (GI) tract with high sensitivities and specificities.[5-9]

In order to achieve such high sensitivities and specificities, the appropriate statistical algorithms must be used to tease out important information from the Raman data. A variety of statistical methods have been developed to classify the tissue as normal or abnormal. For example, many research groups have normalized peak intensities to the four common Raman bands and then performed a student's t-test to identify the peak ratios corresponding to the most significant difference between tissue types.[9, 10] Logistic regression algorithms have also been utilized to distinguish between cancerous and non-cancerous tissue based on Raman spectra. This algorithm was developed by nonlinearly transforming traditional linear regression so the outcome is only 0 (normal) or 1 (cancerous).[11] After normalizing peak ratios, multiple ANOVA have sometimes been used to identify the most diagnostically significant peaks.[12]

Other attempts to analyze data have included using principal component analysis (PCA) to establish differences among and decrease data from Raman spectra.[7, 12, 13] Principal components are a set of virtual spectra; using weighted linear combinations (scores) results in the real, measured spectra with a specified percent-variance. The scores provide information on how the spectra are correlated. Sometimes, this scoring is followed by other statistical analyses such as probabilistic artificial neural networks which can then be used to train the input Raman spectra to correlate with known outputs or pathological categories; this network can then be used to predict the pathology of a new input of Raman spectra. Alternatively, after the spectra are broken down using PCA, linear discriminant analysis (LDA) can be used to maximize differences between pathology groups and minimize differences within groups.[14] Other approaches utilize Fisher discriminant analyses (FDA) to classify the spectra following PCA to search for

nonlinear correlations.[13] The majority of these algorithms undergo cross-validation analysis using the leave-one-out method to assess their validity.[13, 14]

Cluster analysis is one method where similarities between genes are described mathematically, either by measuring the Euclidean distance, angle, or dot products of two *n*-dimensional vectors from series of *n* measurements of genetic information.[15] This algorithm can be similarly applied to determine subtle changes in Raman data. Another process is decision tree learning with genetic algorithms to determine optimal subsets of discriminatory features for pattern recognition.[16, 17] A linear decision binary tree can be used for binary and multi-class, such as Pap smear cell classification.[16] A few drawbacks of these algorithms are that they require a significant amount of time to develop and, once developed, they are only applicable to one type of data set.

The major limitation of these previous applications is that the discrimination algorithms are binary, which are not capable of determining which class the tissue could belong to. Tissue is also not homogenous, there could be multiple tissue types present in a single tissue sample. Therefore, some of these algorithms are run a second and third time to further classify the outcome.[11] More recently, Widjaja, et. al., have combined support vector machines (SVM) with PCA to classify colonic tissues as normal, hyperplastic polyps, or adenocarcinomas[18] However, conventional SVM techniques are used to solve problems with binary solutions. Their modified SVM is able to perform multi-class classification, but still relies on initial binary classification with an incorporated one-against-one strategy to train the model based on probabilities.

Here we present a multi-class approach algorithm based on novel nonlinear statistical methods: Maximum representation and discrimination feature (MRDF)

combined with sparse multinomial logistic regression (SMLR). We will demonstrate this multi-class method in the cervix. Cervical dysplasia is a problem in both the US and throughout the world. Cervical cancer is the second most common malignancy among women worldwide with more than 490,000 cases diagnosed, and 274,000 deaths each year.[19] In the United States alone, it is estimated that in 2008, 3,870 deaths will occur from this disease and 11,070 new cases of invasive cervical cancer will be diagnosed.[20] The mortality rate in the US has been greatly reduced due to effective screening using the Pap smear and effective treatment of precancers (dysplasia).[21] Due to how the disease progresses and regresses (shown in Figure 1), diagnosing the correct grade and progress of cervical dysplasia is very important in treating the disease. Cervical dysplasia is usually classified as one of two groups: 1) Low grade dysplasia which includes human papillomavirus (HPV) and cervical squamous intraepithelial neoplasia (CIN) 1 and 2) High grade dysplasia which includes CIN2, CIN3 and carcinoma in situ (CIS). The progression of disease is shown in figure 1, typically, low grade dysplasia is followed but not treated since approximately, 80% of low grade dysplasia regresses without treatment and less than 1% will develop into cancer.[22] Conversely, 20% of high grade dysplasia will develop into cancer and only 1/3 will regress to a normal state without treatment.[20, 22] Therefore, an algorithm than can differentiate cervical tissue into at least 3 categories is essential: 1) benign cervix, normal, metaplasia and inflammation, 2) low grade and 3) high grade. Metaplasia is often misclassified as dysplastic and therefore an additional category that classifies metaplasia could be beneficial.

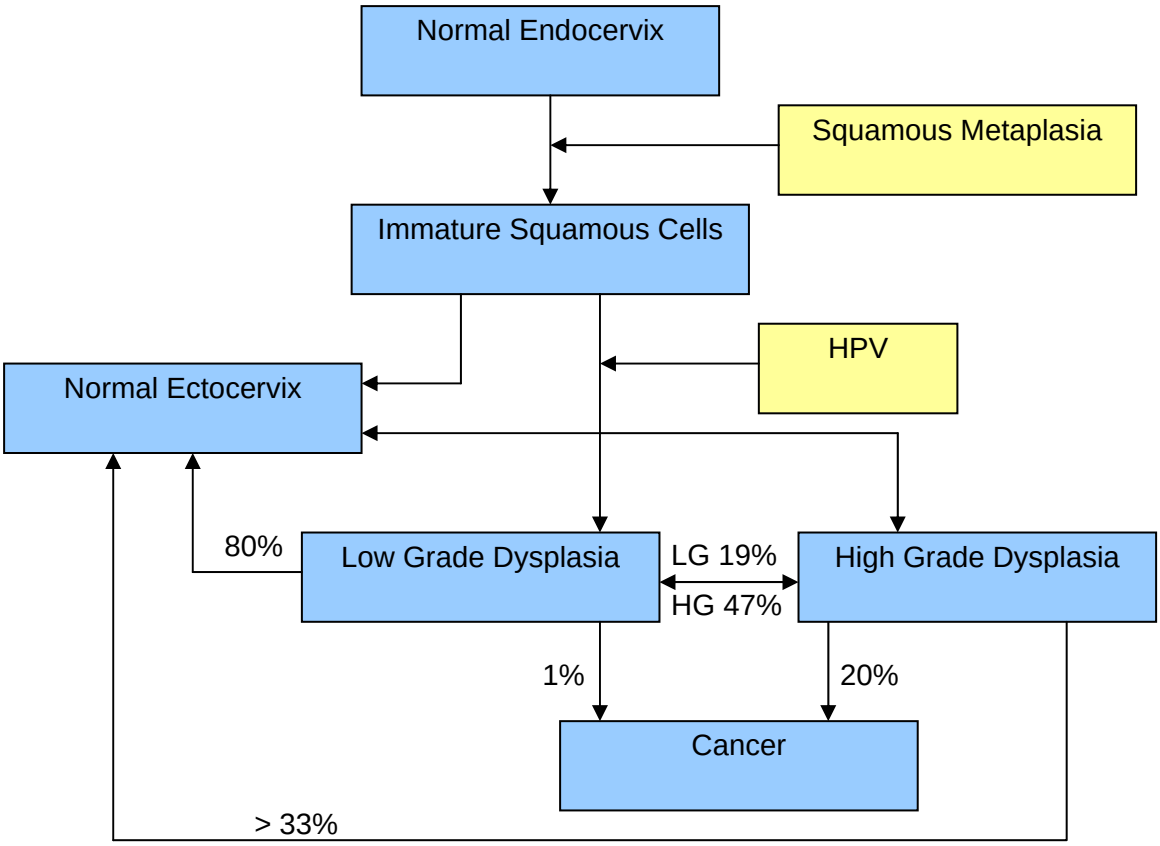


Figure 3.1: A schematic of the progression of normal endocervix cells after squamous metaplasia begins. The cells either transform into normal ectocervix or if infected with HPV may become dysplastic.

In this paper, we demonstrate that by combining Raman spectroscopy data with a more sophisticated statistical method for classification will lead to an enhanced real-time diagnostic tool for cervical dysplasia. First, we will show previous data analyzed with old algorithms. Then, we will establish our new statistical method and use it on the data to show an improvement in specificity and sensitivity. Finally, we will show that we can match or improve classification by Raman versus colposcopy.

3.3 Methods

3.3.1 Data Collection and Instrumentation

A total of 90 patients participated in this study. Measurements were taken from either a procedure that removed diseased cervical tissue or a hysterectomy. The same procedure was followed for data collection regardless of the procedure being performed.

Thirty-three patients undergoing a colposcopy guided biopsy or Loop Electrocautery Excision Procedure (LEEP) were recruited to participate in the study as approved by the Vanderbilt and Copernicus Group Institutional Review Boards (IRBs). Informed consent was obtained from each patient prior to the procedure. The cervix was exposed and visually examined by the doctor. Acetic acid was applied to the cervix to turn abnormal areas white, followed by an application of iodine to clean the tissue and reveal the location of squamous epithelium. Any abnormal tissue was removed and histopathology was performed. Raman spectra were acquired after the application of acetic acid but before the application of iodine and the removal of tissue. Spectra were measured from each visually abnormal area (1-6 measurements) and one visually normal area. The patient's age, date of last period, abnormal Pap smear result and menopausal status were noted upon chart review.

Additionally, 33 patients undergoing hysterectomy were recruited to participate in the study as approved by Vanderbilt IRB. Informed consent was obtained from each patient prior to the procedure. The cervix was then exposed and visually examined by the doctor. Acetic acid was applied to the cervix to keep the procedure similar to that of the dysplasia patients. If the cervix was visually normal, spectra were measured from

multiple normal areas of tissue. Measured areas were marked, the hysterectomy then proceeded as required and the removed tissue was histopathology was performed.

Raman spectra were acquired using a portable Raman spectroscopy system consisting of a 785 nm diode laser (Process Instruments, Inc., Salt Lake City, UT), 7 (300um) around 1 (400um) beam-steered fiber optic probe (Visionex Inc.), imaging spectrograph (Kaiser Optical Systems, Inc., Ann Arbor, MI), and back-illuminated, deep-depletion, charge coupled device (CCD) camera (Princeton Instruments, Princeton, NJ), all controlled with a laptop computer. For this study, the fiber optic probe delivered 80 mW of incident light onto the tissue and collected the scattered light for 5 seconds. In all cases, the overhead fluorescent lights and colposcope light were turned off during the measurements. Any luminescent lights were left on but turned away from the measurement site.

3.3.2 Data Processing

The wavenumber axis was calibrated using neon-argon lamp, acetaminophen, and naphthalene standards each day. The signal from the Raman spectrum was binned along the vertical axis to create a single spectrum per measurement site. Prior to any signal processing, the spectrum was truncated to only include the region from about 990 cm^{-1} to 1850 cm^{-1} to eliminate the Raman peaks due to the silica present in the fiber optic probe. The spectrum was then binned along the wavenumber axis in 3.5 cm^{-1} intervals and noised-smoothed with a 2^{nd} order Savitzky-Golay filter. Additionally, fluorescence background was removed using an automated, modified polynomial fitting method that utilizes a 5^{th} degree polynomial to fit the fluorescence baseline.[23] Once noise

smoothing, fluorescence subtraction were done spectra were normalized to its mean spectral intensity across all Raman bands and were used for subsequent data analysis.

3.3.3 Statistical Analysis

To compare and contrast the binary versus multinomial class techniques, two different algorithms have been developed to classify cervical data. The first is a binary algorithm that is based on peak rations and logistic regression. The second is a multi-class probabilistic algorithm that is based on machine support vectors and nonlinear logistic regression. Both algorithms are described in detail below.

3.3.4 Statistical Analysis – Binary

The first step in using Raman spectra is to develop a basic algorithm to discriminate between abnormal and normal tissue. First, the mean and standard deviation at each wavenumber of the spectra within each pathology group was calculated to characterize the overall spectral trends for each group. A Student's t-test was performed at each wavenumber between individual pairs of pathology groups to identify regions of spectral distinction between two different pathologies. Any major peak that showed statistical differences at the level of $p < 0.01$ between normal ectocervix spectra and high grade dysplasia spectra was chosen as an input for the algorithm. Thus, the inputs to the algorithm are the normalized intensity values at 1006, 1055, 1244, 1305, 1324, 1450, 1550, 1657 cm^{-1}. The classification model was constructed to automatically classify spectra into one of two categories (high grade dysplasia or benign cervix) using a two-tiered logistic regression model.[11] The first algorithm was developed to distinguish

normal from all other pathologies (metaplasia, high grade dysplasia) and the second algorithm discriminated high grade dysplasia from other pathologies (metaplasia).

The first algorithm was trained using a training set to classify a spectrum as either normal ectocervix (score=0) or high grade dysplasia (score =1); the algorithm was then tested using a separate validation set. The training and validation sets were randomly generated by dividing the normal ectocervix and high grade dysplasia data sets into a training set (two-thirds of the patients) and a validation set (one-third of the patients). The training and validation sets were divided by patients, not individual spectra, such that all spectra from one patient were either in the training set or the validation set, but not both.

The algorithm then output a score, which represents the likelihood that the input data represents high grade dysplasia. Data from the squamous metaplasia were also included as part of the validation set for the model even though no data from this category were included in the training set to examine the possibility that a single algorithm could discriminate all spectra of benign pathology from dysplasia spectra (see discussion). Since the specificity of this single algorithm model was less than desired due primarily to misclassifications of squamous metaplasia spectra, a second logistic regression algorithm was developed to separate high grade dysplasia from squamous metaplasia to increase the specificity of the overall model.

Any spectra from the test set with a score greater than 0.5 from the first algorithm formed the test set into the second algorithm (thus there were 8 high grade dysplasia, 8 squamous metaplasia, and 4 normal ectocervix spectra). The training set for the second algorithm was formed using only the high grade dysplasia spectra (29 spectra, score=1)

and squamous metaplasia (29 spectra, score=0) spectra as there were not enough spectra to create separate training and validation sets. The same Raman bands from the first algorithm were also used as inputs in the second algorithm, but the output was a value (score) that represents the probability that the spectra were measured from an area of high grade dysplasia as compared with squamous metaplasia. While the data did classify well using these algorithms, it was clear that we were losing some information in the Raman spectra by only looking at binary classification.

3.3.5 Statistical Analysis – Multi-class

Maximum representation and discrimination feature (MRDF) combined with sparse multinomial logistic regression (SMLR), was used to develop a multi-class diagnostic algorithm.[24] This algorithm is a two step process; (1) extraction of diagnostic features from spectra using nonlinear MRDF and (2) classification based on these nonlinear features into corresponding tissue categories using SMLR. Figure 3.2 shows a flow chart of this algorithm.

MRDF is a method of feature extraction; it maximally extracts the diagnostic information otherwise hidden in a set of measured spectral data by reducing its dimensionality through a set of mathematical transforms as shown in figure 3.2.b. Given a set of input data comprising of spectra from different classes with a given dimensionality, nonlinear MRDF determines a set of nonlinear transformations of the input data that optimally discriminates between the different classes in a reduced dimensionality space. It invokes nonlinear transforms (restricted order polynomial mappings of the input data) in two successive stages. In the first stage, the input spectral

data $\mathbf{x} = [x_1, x_2, ...x_N]^T$ (intensities corresponding to Raman shifts of the spectra) from each tissue type are raised to the power p' to produce the associated nonlinear input vectors $\mathbf{x}_{p'} = [x_1^{p'}, x_2^{p'}, ..., x_N^{p'}]$. These vectors are then subject to a transform Φ'_M such that $\mathbf{y}'_M = \Phi'^T_M \mathbf{x}_{p'}$ and are the first stage output features in the nonlinear space of reduced dimension $M << N$. In the second stage, the reduced M dimensional output features $\mathbf{y}'_M$ for each tissue type are further transformed nonlinearly to the power p to produce higher order features $\mathbf{y}'_{Mp} = [y'^p_1, y'^p_2, ..., y'^p_M]$, and a second transform Φ_K is computed so as to yield the final output features $\mathbf{y}_K = \Phi^T_K \mathbf{y}_{Mp}$ in the nonlinear space of dimension K ($K \leq M$).[25]

SMLR is a method of supervised classification. It is a probabilistic multi-class model based on sparse Bayesian machine-learning framework of statistical pattern recognition. The central idea of SMLR is to separate a set of labeled input data into its constituent classes by predicting the posterior probabilities of their class-membership. It computes the posterior probabilities (from the equations shown in figure 3.2.b) using a multinomial logistic regression model and constructs a decision boundary that separates the data into its constituent classes based on the computed posterior probabilities following Bayes' rule (i.e. a class is assigned to a spectra for which its posterior probability is the highest).[25] Traditional statistical methods have focused primarily on using binary classification. However, this method is limited when looking at complicated diseases, like stages of cancer. A novel, multi-class method is more suitable for such applications.

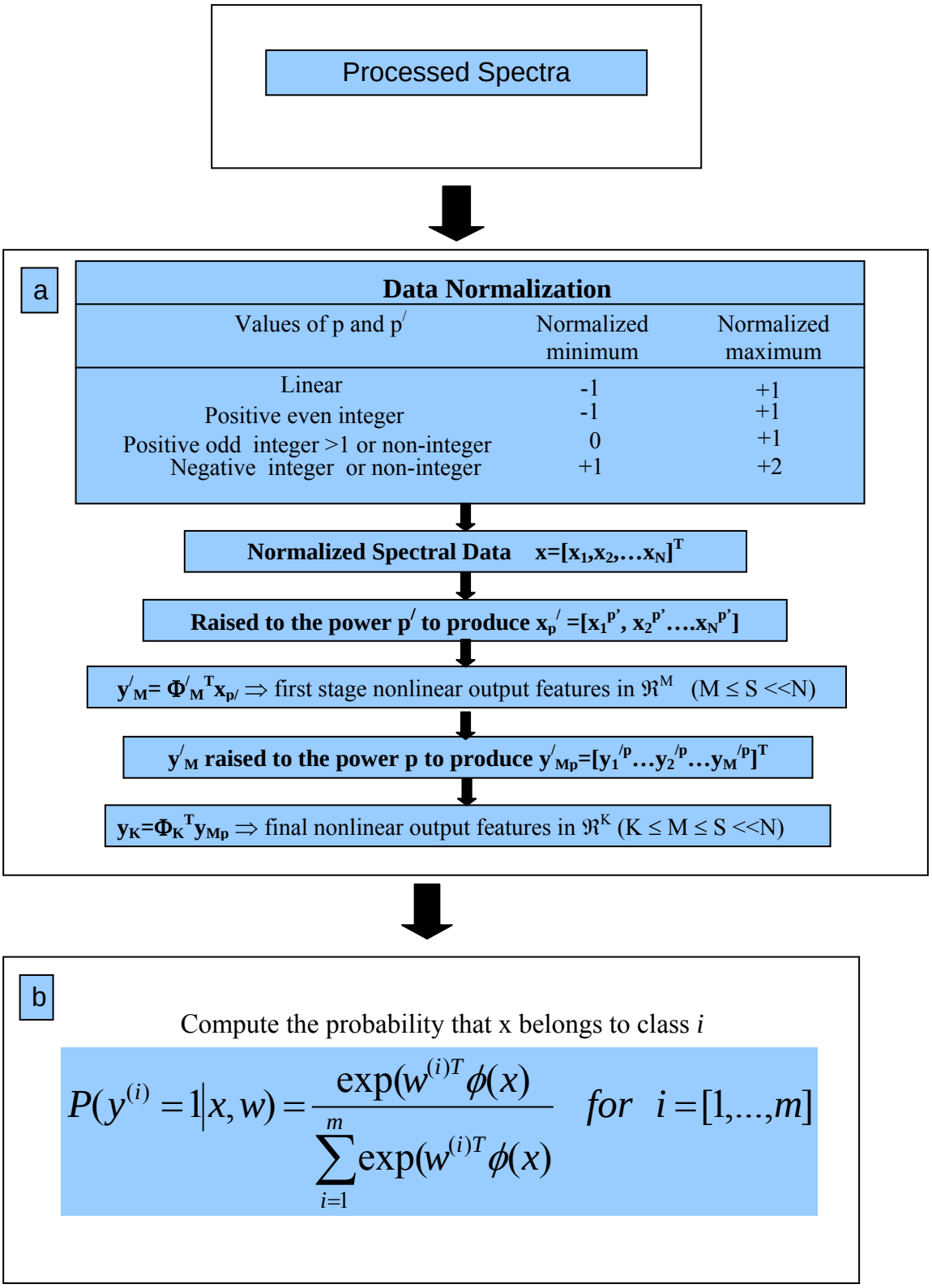


Figure 3.2: Flow chart of the multi-class discrimination algorithm.

3.4 Results

Using both algorithms, a total of 29 high grade dysplasia (from 19 patients), 6 low grade dysplasia (from 5 patients), 29 squamous metaplasia (from 20 patients) and 100 normal ectocervix (from 47 patients) were classified to compare the sensitivity and specificity of the two algorithms.

First, the resulting spectra were then correlated with the corresponding histopathologic diagnosis to characterize the differences between various diagnostic categories. Figure 3.3 shows the mean spectra for the full data set for each of the different categories.

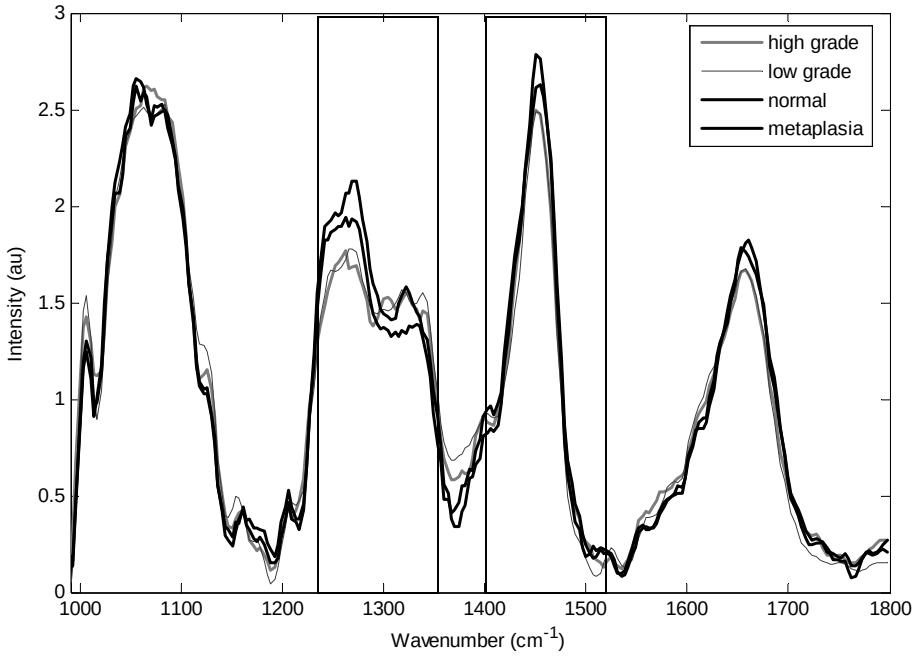


Figure 3.3: Average Raman spectra for normal ectocervix, low grade dysplasia, high grade dysplasia and metaplasia, the boxed regions are regions that are different.

Peaks were found at 1006, 1058, 1086, 1244, 1270, 1324, 1450, 1550, 1655 cm^{-1} in most spectra. Although the peak shapes and locations are consistent across all pathology classifications, there are small but significant differences in peak intensities among the different pathology categories. Several spectral regions show statistically significant differences in comparing precancer to the normal ectocervix. For example, in low grade spectra the 1324 cm^{-1} peak increases as compared with normal ectocervix, similar to the high grade precancer/normal ectocervix spectral comparison. But, the intensity of the 1272 cm^{-1} and 1450 cm^{-1} peaks in low grade precancer spectra seems to remain similar to that seen in normal ectocervix, unlike in the high grade precancer spectra. These differences are very subtle so there is a need for statistical approaches.

A binary algorithm based on LDA was applied, using the output of this algorithm; we were able to distinguish between high grade precancer and benign areas of the cervix (normal ectocervix and squamous metaplasia) with sensitivity of 89% and specificity of 88%. Due to insufficient numbers (only data from 7 patients with evidence of LGSIL), we did not include low-grade data in this analysis. The limitation of this particular discrimination algorithm is that it is binary, which does not allow for the multiple classes that the tissue could belong to. In addition, the inputs for the algorithm were selected as normalized peak intensities thus ignoring other potentially useful information present in the spectra. To address these limitations, a second discrimination algorithm was developed based on novel nonlinear statistical methods: MRDF combined with SMLR.

In order to determine the effectiveness of a new discrimination algorithm as compared to the one used previously, we used the same 66 patients as before and classified this data using an algorithm based on MRDF and SMLR. The diagnostic

algorithm using this method has the potential to discriminate using the complete *in vivo* Raman spectra acquired from the human cervix, simultaneously into the different pathological categories. Unbiased performance estimates were obtained using leave-one-patient-out cross validation. The results indicate that Raman spectroscopy can distinguish high grade precancer from normal ectocervix and squamous metaplasia with a higher sensitivity and specificity than the binary algorithm (sensitivity - 92% (probability of disease classified correctly (LGSIL and HGSIL) and specificity - 96% probability of benign being classified correctly (normal and metaplasia) as shown in table 3.1.

Table 3.1. Classification using algorithm based on MRDF and SMLR.

Classification	**Pathology based classification**			
Raman Algorithm based classification	**High Grade**	**Low Grade**	**Metaplasia**	**Normal**
High Grade	**20**	0	0	0
Low Grade	0	**5**	0	0
Metaplasia	0	0	**20**	3
Normal	1	2	1	**66**

High grade spectra were classified correctly 95% of the time, and only one misclassified as normal. Low grade data was never classified as high grade and was misclassified as normal 29% of the time. Very few low grade spectra were including in the analysis but this method now has a similar sensitivity and much higher specificity

than colposcopy guided biopsy in expert hands (sensitivity of 87% and specificity of 72%).

In order to ensure that our algorithm applies also to low-grade cervical precancers, we added twenty-seven patients which increased the number of samples within each category with an emphasis on low grade lesions. Raman spectra from a total of 93 patients were analyzed using the algorithm based on MRDF and SMLR with leave-one-patient-out cross validation. The performance of the model is reported in table 3.2.

Table 3.2. Classification using algorithm based on MRDF and SMLR.

Classification	**Pathology based classification**			
Raman Algorithm based classification	**High Grade**	**Low Grade**	**Metaplasia**	**Normal**
High Grade	**24**	1	1	3
Low Grade	0	**18**	0	3
Metaplasia	0	0	**19**	10
Normal	5	3	10	**208**

The result of SMLR is a set of predictive values (or posterior probabilities) and these were obtained by using leave-one-patient-out cross-validation. Figure 3.4 plots the predictive posterior probabilities of being classified as high grade dysplasia, low grade dysplasia, squamous metaplasia, and normal ectocervix for the normalized Raman spectra of the corresponding cervical tissue sites. Even though emphasis was placed on

collecting low grade spectra, only 23 low grade spectra from 19 patients are represented in this study which may be a reason for the higher misclassification rate. Even though we had some misclassifications with the low-grade data, overall, more than 88% of the data from this set classified correctly. More clinical data from low grade cervix is needed to determine the full capability of this algorithm to differentiate between normal and low grade.

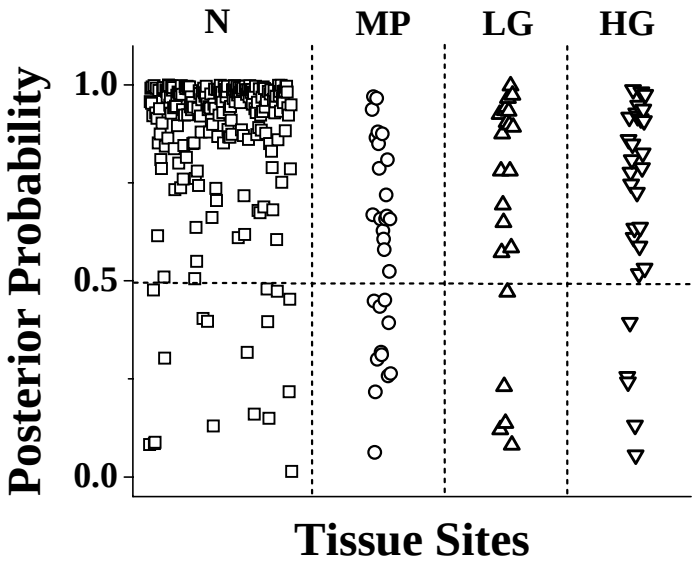


Figure 3.4: Posterior probabilities of being classified as normal ectocervix (N), low grade (LG) dysplasia, high grade dysplasia (HG) and metaplasia (MP).

3.5 Discussion

Raman spectroscopy has the power to optically identify subtle changes in tissue that can lead to diseases like cancer. Many statistical methods have been developed to tease out important, clinical parts of Raman spectra, allowing them to be correlated to specific pathological conditions. Although binary methods have traditionally been used

to classify spectroscopy data, a more sophisticated method that is able to classify multiple classes at the same time is necessary as Raman spectroscopy moves closer to the clinic. Raman spectroscopy combined with a multi-class discrimination algorithm has great potential for biological applications, especially within tissue. This paper demonstrates that when using the multi-class algorithm, we can improve the sensitivity from 92% to 98% and the specificity from 81% to 96%. Both methods (binary and multi-class) are an improvement over the current method of diagnosis, colposcopy guided biopsy in expert hands has a sensitivity of 87% and a specificity of 72% showing the capabilities of Raman spectroscopy.

One major concern when taking *in vivo* measurements is that a certain sample volume may have two different pathological classifications. The sample could be 75% metaplasia and 25% high grade dysplasia. Therefore, in a binary algorithm that separates between high grade and metaplasia would classify this sample as metaplasia since the sample is dominated by metaplasia. But with the multi-class algorithm, we may be able to show that this sample is mostly metaplasia, but has spectral contributions from the high grade dysplastic tissue. This feature would prevent misdiagnosing the tissue as normal instead of metaplasia with some high-grade regions. This is an important benefit of using the multi-class algorithm, even though the implementation of it is more complicated than a binary algorithm.

The goal of this present study was to develop a multivariate statistical algorithm capable of simultaneously classifying Raman spectral data acquired *in vivo* from human cervical tissues into high grade dysplasia, low grade dysplasia, squamous metaplasia, and normal ectocervix. The first task for the development of such an algorithm is the

extraction of diagnostically relevant features from the observed spectra by reducing the dimensionality of the measured spectral variables. For good classification performance, the extracted features should contain sufficient class-discriminatory information. Most of the published reports on spectroscopic diagnostic algorithms have reported using standard linear techniques like PCA and FDA to extract diagnostic features from the measured spectra of tissue.[26-29] Although these linear techniques have the advantage of providing closed-form solutions, which make them relatively easy to implement, they are limited because they only extract information from the second-order correlation in the data and ignore higher order correlations that could be useful for improved discrimination. Use of nonlinear techniques is required for this purpose.[30] There are several nonlinear methods that exist for feature extraction in the pattern recognition literature; most of them are iterative and often need a priori selection of a number of parameters associated with the learning or the optimization technique used.[30] They also are limited by problems with convergence.[30] One major advantage of the nonlinear MRDF technique is that unlike the iterative nature of other nonlinear feature extraction algorithms, it provides a closed-form expression of the nonlinear transform for maximum discrimination.[31, 32] Another advantage of using this method to classify spectral data is that it has the ability to separate classes that are not linearly separable. As spectral data tends to be non-symmetric, using MRDF can lead to spectral separations with higher accuracy.

This increased sensitivity makes Raman spectroscopy superior over other types of spectroscopy and therefore ideal for detection of small changes in early dysplasia. Although the number of low grade spectra in this study is few, we could identify the low

grade spectra 74% of the time. Once diagnostic features are extracted from the measured spectral data, the final task of the algorithm is to classify these extracted features into respective tissue categories. The major advantage of using the SMLR approach for classification is that since it is based on a Bayesian framework, it is able to predict the posterior probability of class-membership of the investigated tissue site. This idea is demonstrated in Figure 3.4 where the predicted posterior probabilities of the different cervical tissue sites classified as high grade dysplasia, low grade dysplasia, squamous metaplasia, and normal ectocervix are plotted. One may also note that most of the dysplastic sites have been classified with a posterior probability of greater than 80% into the corresponding tissue categories. The probabilistic approach can offer an important advantage by making it possible to further interrogate these sites, especially when the goal is to correctly identify all abnormal sites for accurate screening of cervical dysplasia. An additional advantage of the new algorithm is that it provides the posterior probability of samples belonging to the different diagnostic categories. We expect this to be extremely useful in a clinical setting because health providers could recheck samples having lower posterior probability of belonging to one category by using a traditional biopsy method.

Other groups have suggested that optical technologies are capable of distinguishing high grade dysplasia or cancer from normal cervix, but have had little success at differentiating low grade dysplasia from normal or high grade. In this paper, we have demonstrated that by using MRDF with SMLR, Raman spectroscopy is capable of picking out some of the subtle changes that occur during the early stages of dysplasia. Unfortunately, we were unable to collect a large number of low grade dysplasia data and

future studies need to be focused on low grade data collection. Also, these algorithms need to be usable in real-time clinical settings which will also be developed in the future.

3.6 Conclusions

Using a probability-based robust diagnostic algorithm capable of simultaneously discriminating *in vivo* Raman spectra, acquired from human cervical tissues, into various pathological categories improves performance compared to more traditional methods by allowing for multi-class discrimination. The results indicate that Raman spectroscopy in conjunction with the diagnostic algorithm can distinguish dysplasia from normal ectocervix (including metaplasia) with a classification accuracy of 95%. One additional advantage of the algorithm developed in this study is that it provides the posterior probability of samples belonging to the different diagnostic categories. This is expected to be extremely useful in a clinical setting, because clinicians could recheck any sample having a lower posterior probability of belonging to one category with the conventional biopsy method. These discrimination techniques are not only applicable to the cervix, but could be used in the spectral data analysis of tissue types that require a multi-class diagnosis, like GI and skin cancers.

3.7 Acknowledgements

The authors would like to acknowledge the financial support of the NCI/NIH (R01-CA95405). We would also like to thank the doctors and staff at Vanderbilt University and at Tri-state women's health for all of their help.

3.8 References

1. A. Mahadevan-Jansen and R. Richards-Kortum, "Raman Spectroscopy for the detection of cancers and precancers," *J Biomed Opt*, 1(1), 31-70, 1996.

2. M. S. Feld, R. Manoharan, J. Salenius, J. Orenstein-Carndona, T. J. Roemer, J. F. Brennan III, R. R. Dasari and Y. Wang, "Detection and characterization of human tissue lesions with near-infrared Raman spectroscopy", *Advances in Fluorescence Sensing Technology II,* SPIE, 2388(99-104), 1995.

3. E. B. Hanlon, R. Manoharan, T. W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari and M. S. Feld, "Prospects for in vivo Raman spectroscopy," *Physics in Medicine and Biology*, 45(2), R1-R59, 2000.

4. U. Utzinger and R. R. Richards-Kortum, "Fiber optic probes for biomedical optical spectroscopy," *J Biomed Opt*, 8(1), 121-47, 2003.

5. M. V. P. Chowdary, K. K. Kumar, J. Kurien, S. Mathew and C. M. Krishna, "Discrimination of normal, benign, and malignant breast tissues by Raman spectroscopy," *Biopolymers*, 83(5), 556-569, 2006.

6. T. R. Hata, T. A. Scholz, I. V. Ermakov, R. W. McClane, F. Khachik, W. Gellermann and L. K. Pershing, "Non-invasive Raman spectroscopic detection of carotenoids in human skin," *Journal of Investigative Dermatology*, 115(3), 441-448, 2000.

7. M. G. Shim, L. Song, N. E. Marcon and B. C. Wilson, "In vivo near-infrared Raman spectroscopy: Demonstration of feasibility during clinical gastrointestinal endoscopy," *Photochemistry and Photobiology*, 72(1), 146-150, 2000.

8. A. P. Oliveira, R. A. Bitar, L. Silveira, R. A. Zangaro and A. A. Martin, "Near-infrared Raman spectroscopy for oral carcinoma diagnosis," *Photomedicine and Laser Surgery*, 24(3), 348-353, 2006.

9. A. Mahadevan-Jansen, W. F. Mitchell, N. Ramanujam, U. Utzinger and R. Richards-Kortum, "Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue in vivo," *Photochemistry and Photobiology*, 68(3), 427-431, 1998.

10. U. Utzinger, D. L. Heintzelman, A. Mahadevan-Jansen, A. Malpica, M. Follen and R. Richards-Kortum, "Near-infrared Raman spectroscopy for in vivo detection of cervical precancers," *Applied Spectroscopy*, 55(8), 955-959, 2001.

11. A. Viehoever Robichaux, E. Kanter, H. Shappell, D. Billheimer, H. Jones and A. Mahadevan-Jansen, "Characterization of Raman Spectra Measured In vivo for the Detection of Cervical Dysplasia," *Applied Spectroscopy*, 61(9), 986-993, 2007.

12. N. Stone, P. Stavroulaki, C. Kendall, M. Birchall and H. Barr, "Raman spectroscopy for early detection of laryngeal malignancy: preliminary results," *Laryngoscope*, 110(10 Pt 1), 1756-63, 2000.

13. A. Mahadevan-Jansen, M. F. Mitchell, N. Ramanujam, A. Malpica, S. Thomsen, U. Utzinger and R. Richards-Kortum, "Near-infrared Raman spectroscopy for in vitro detection of cervical precancers," *Photochem Photobiol*, 68(1), 123-32, 1998.

14. C. Kendall, N. Stone, N. Shepherd, K. Geboes, B. Warren, R. Bennett and H. Barr, "Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus," *J Pathol*, 200(5), 602-9, 2003.

15. M. B. Eisen, P. T. Spellman, P. O. Brown and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proc Natl Acad Sci U S A*, 95(25), 14863-8, 1998.

16. B. B. Chai, T. Huang, X. H. Zhuang, Y. X. Zhao and J. Sklansky, "Piecewise linear classifiers using binary tree structure and genetic algorithm," *Pattern Recognition*, 29(11), 1905-1917, 1996.

17. G. Zhou, Y. Chen, Z. Wang and H. Song, "Genetic local search algorithm for optimization design of diffractive optical elements," *Appl Opt*, 38(20), 4281-90, 1999.

18. E. Widjaja, W. Zheng and Z. Huang, "Classification of colonic tissues using near-infrared Raman spectroscopy and support vector machines," *Int J Oncol*, 32(3), 653-62, 2008.

19. G. P. Parham, V. V. Sahasrabuddhe, M. H. Mwanahamuntu, B. E. Shepherd, M. L. Hicks, E. M. Stringer and S. H. Vermund, "Prevalence and predictors of squamous intraepithelial lesions of the cervix in HIV-infected women in Lusaka, Zambia," *Gynecologic Oncology*, 103(3), 1017-1022, 2006.

20. "American Cancer Society," 2008.

21. E. R. Myers, D. C. McCrory, S. Subramanian, N. McCall, K. Nanda, S. Datta and D. B. Matchar, "Setting the target for a better cervical screening test: characteristics of a cost-effective test for cervical neoplasia screening," *Obstet Gynecol*, 96(5 Pt 1), 645-52, 2000.

22. "American Medical Association," 1999.

23. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Applied Spectroscopy*, 57(11), 1363-1367, 2003.

24. S. K. Majumder, S. Gebhart, M. D. Johnson, R. Thompson, W. C. Lin and A. Mahadevan-Jansen, "A probability-based spectroscopic diagnostic algorithm for simultaneous discrimination of brain tumor and tumor margins from normal brain tissue," *Appl Spectrosc*, 61(5), 548-57, 2007.

25. S. K. Majumder, E. Kanter, A. R. Viehoever, H. Jones and A. Mahadevan-Jansen, "Near-infrared Raman spectroscopy for in-vivo diagnosis of cervical dysplasia: a probability-based multi-class diagnostic algorithm", *SPIE,* pp 64300Q-64311), 2007.

26. N. Ramanujam, M. F. Mitchell, A. Mahadevan-Jansen, S. L. Thomsen, G. Staerkel, A. Malpica, T. Wright, N. Atkinson and R. Richards-Kortum, "Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths," *Photochem Photobiol*, 64(4), 720-35, 1996.

27. S. K. Majumder, S. K. Mohanty, N. Ghosh, P. K. Gupta, D. K. Jain and F. Khan, "A pilot study on the use of autofluorescence spectroscopy for diagnosis of the cancer of human oral cavity," *Current Science*, 79(8), 1089-1094, 2000.

28. C. Y. Wang, C. T. Chen, C. P. Chiang, S. T. Young, S. N. Chow and H. K. Chiang, "A probability-based multivariate statistical algorithm for autofluorescence spectroscopic identification of oral carcinogenesis," *Photochem Photobiol*, 69(4), 471-7, 1999.

29. E. N. Atkinson, M. F. Mitchell, N. Ramanujam and R. Richards-Kortum, "Statistical techniques for diagnosing CIN using fluorescence spectroscopy: SVD and CART," *J Cell Biochem Suppl*, 23(125-30, 1995.

30. A. K. Jain, R. P. W. Duin and J. C. Mao, "Statistical pattern recognition: A review," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4-37, 2000.

31. A. Talukder, "Nonlinear feature extraction for pattern recognition applications," PhD Thesis Carnegie Mellon University, 1999.

32. A. Talukder and D. Casasent, "General methodology for simultaneous representation and discrimination of multiple object classes," *Optical Engineering*, 37(3), 904-913, 1998.

CHAPTER IV

VARIATIONS IN THE NORMAL CERVIX

4.1 Introduction

Cervical cancer is the second most common malignancy among women worldwide, with more than 490,000 cases diagnosed and 274,000 deaths each year. [1] In the United States alone, it is estimated that in 2008, 3,870 deaths will occur from this disease, and 11,070 new cases of invasive cervical cancer will be diagnosed. [2] Several recent meta-analyses have reported low Pap smear sensitivities in the range of 50 percent, but as low as 20 percent. [3, 4] Most of these studies, however, indicate that the Pap smear is generally a very specific test, meaning that cytology correctly identifies a high proportion of women who do not have high grade lesions or cancer. Although cervical cancer does affect young women, many older women do not realize that the risk of developing cervical cancer is still present as they age. Slightly over 20% of women with cervical cancer are diagnosed when they are over 65. [5] Further, cervical cancer in Hispanic women occurs at a rate that is more than two times that of non-Hispanic white women. African-American women develop this cancer about 50% more often than non-Hispanic white women. [5] Thus there is a continued need for an effective diagnostic and guidance tool.

Optical methods, including fluorescence spectroscopy, reflectance spectroscopy, optical coherence tomography (OCT), and Raman spectroscopy have been investigated as s potential new diagnostic tool. In particular, Raman spectroscopy has shown

considerable promise as a new technology to detect and screen for cervical precancers with improved specificity over traditional methods. Current studies suggest that Raman spectroscopy can differentiate between normal cervical tissue, metaplasia, low grade squamous intraepithelial lesion (LGSIL), and high grade squamous intraepithelial lesion (HGSIL) with an accuracy of 90%. The most difficult pathology to classify is LGSIL, which is often misclassified as normal (Chapter 3). One of the goals of this study, then, was to improve the ability to differentiate between LGSIL and healthy squamous epithelium.

One method to improve the differentiation of LGSIL from healthy tissue is to determine the source of spectral variability in the normal cervix. Many different physiological factors may affect the normal cervix, such as hormonal status and previous disease of the cervix. Several studies have shown that menopausal state and location in the menstrual cycle impact the optical properties. For example, a study done by Gill *et. al*. showed that there was a statistical difference between pre- and post-menopausal fluorescence signal, and that the post-menopausal women had a higher average fluorescence signal. This difference is likely due to the changes in collagen cross-linking of the cervix that occur during menopause. [6] Another study done by Cox *et. al*. looked at the effects of the menstrual cycle on fluorescence measurements. It indicated that the primary spectral differences seen relative to the menstrual cycle were caused by hemoglobin absorption, but this variability can be avoided if measurements are not taken in the first eight days of the cycle. [7] The above studies used only fluorescence spectroscopy, but other changes in the cervix may be detected by Raman since the biochemical composition of the cervix changes during the menstrual cycle.

Variations in normal cervical spectra due to menopausal status and time point in the menstrual cycle, previous vaginal deliveries and evidence of previous disease were investigated in this study. These differences were then used to improve the ability of Raman spectroscopy to diagnose low grade dysplasia. Raman spectra from "normal patients" (no evidence of cervical dysplasia) and where then analyzed using advanced statistical methods. Based on the results from investigating the variations in the normal cervix, more stratifications, in particular location in the menstrual cycle where included in the diagnostic algorithm and the ability to classify LGSIL improves to 97%.

4.2 Methods

4.2.1 Clinical Study Design-Pap Smear Patients

Ninety-one patients undergoing a routine annual Pap smear were recruited to participate in the study as approved by the Institutional Review Board (IRB). To be eligible for enrollment, the patient must be undergoing a routine Pap smear, be between the ages of 18-75, and have a cervix (no history of a hysterectomy). Informed consent was obtained from each patient prior to the procedure. The cervix was exposed, visually examined by the attending physician, and wiped clean with a dry cotton swab and then with a saline solution. Multiple Raman spectra (3-5) of normal appearing sites were measured *in vivo,* and tissue sites were recorded as squamous epithelium, columnar epithelium, or at the squamous-columnar junction as determined by the attending physician. The Pap procedure then proceeded according to standard clinical protocol. The acquired spectra were considered normal if the Pap smear came out negative. The

patient's age, last menstrual period, artificial hormones (including contraception), and menopausal status were all noted upon chart review.

4.2.2 Clinical Study Design- Colposcopy Patients

Thirty one patients undergoing colposcopy-guided biopsy were recruited to participate in the study as approved by the Vanderbilt and Copernicus Group IRBs. To be eligible for enrollment, the patient must be undergoing a colposcopy-guided biopsy, be between the ages of 18-75, and have a cervix (no history of a hysterectomy). Informed consent was obtained from each patient prior to the procedure. The cervix was exposed and visually examined by the attending physician. Acetic acid was applied to the cervix to turn abnormal areas white, enabling visualization of abnormal areas, and iodine was applied to clean the tissue and show the location of squamous epithelium. The abnormal tissue was removed and the pathology was examined. Spectra were measured after the application of the acetic acid and before the application of the iodine (if needed). Spectra were acquired from each area where a biopsy was taken and one visually normal area. The patient's age, last menstrual period, abnormal Pap smear result, and menopausal status were all noted upon chart review.

4.2.3 Data Collection

Raman spectra were collected from multiple sites *in vivo* using a portable Raman spectroscopy system consisting of a 785 nm diode laser (Process Instruments, Inc., Salt Lake City, UT), beam-steered fiber optic probe (Visionex Inc., Atlanta, GA), imaging spectrograph (Kaiser Optical Systems, Inc., Ann Arbor, MI), and back-illuminated, deep-

depletion, thermo-electrically cooled charge coupled device (CCD) camera (Roper Scientific, Inc., Princeton, NJ), all controlled with a laptop computer. A photograph of the system can be seen in figure 4.1; more details of the system can be found elsewhere [8]. The fiber optic probe deliverers 80mW of incident light onto the tissue for 3 seconds. For each measurement, the overhead fluorescent lights were turned off.

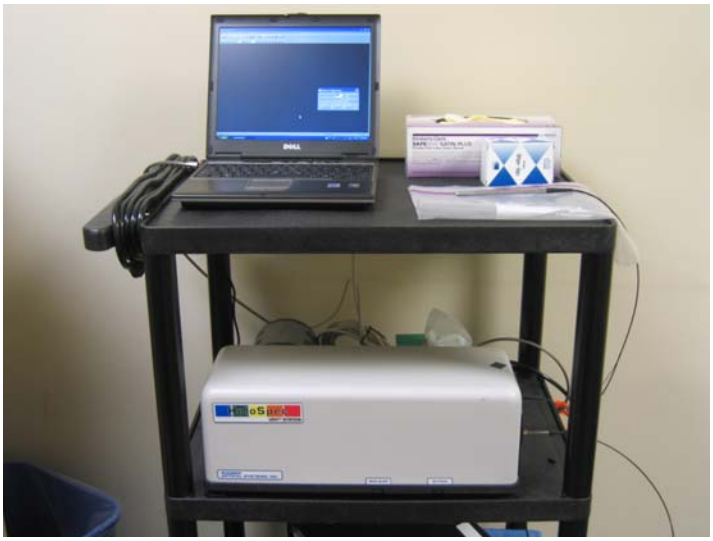


Figure 4.1: Photograph of the system

Spectral calibration of the system was preformed each day using a neon-argon lamp, naphthalene and acetaminophen standards to correct for system wavenumber, laser excitation, and throughput variations. The spectra were further corrected using a calibrated tungsten lamp to correct for system variations. The spectra were processed for fluorescence subtraction and noise smoothing using a modified mean polynomial fitting approach [8]. Following data processing, each spectrum was normalized to its mean spectral intensity across all Raman bands to account for overall intensity variability.

These normalized spectra were categorized according to histology (squamous epithelium, columnar epithelium, or at the junction) as determined by the attending physician, and by menopausal status and used for further comparison and analysis.

4.2.4 Statistical Analysis

The analysis technique that is used in this paper has been described elsewhere[9]. The process consists of two steps - the first is extraction of diagnostic features from the spectra using the nonlinear maximum representation and discrimination feature (MRDF). The processed data set undergoes a two-step, non-linear transform to extract relevant features that provide the best class separation. The second step is a probabilistic classification scheme based on sparse linear multinomial logistic regression (SMLR) for classifying the nonlinear features into corresponding tissue categories. SMLR is a probabilistic multi-class model based on a Bayesian machine-learning framework of statistical pattern recognition. The main focus of SMLR is to separate a set of labeled input data into its class by predicting the posterior probabilities of their class membership. All classification was done using leave-one-patient-out cross-validation.

4.3 Results

4.3.1. Variations in Hormonal Status

The spectra in Figure 4.2 are shown stratified by menopausal status. The data is separated into one of 4 groups: pre-menopausal proliferative phase (days 1-14 of the menstrual cycle), or pre-menopausal before ovulations (PBO); pre-menopausal secretory

phase (days 15-28+ of the menstrual cycle), or pre-menopausal after ovulations (PAO); peri-menopausal (PERI); and post-menopausal (POST).

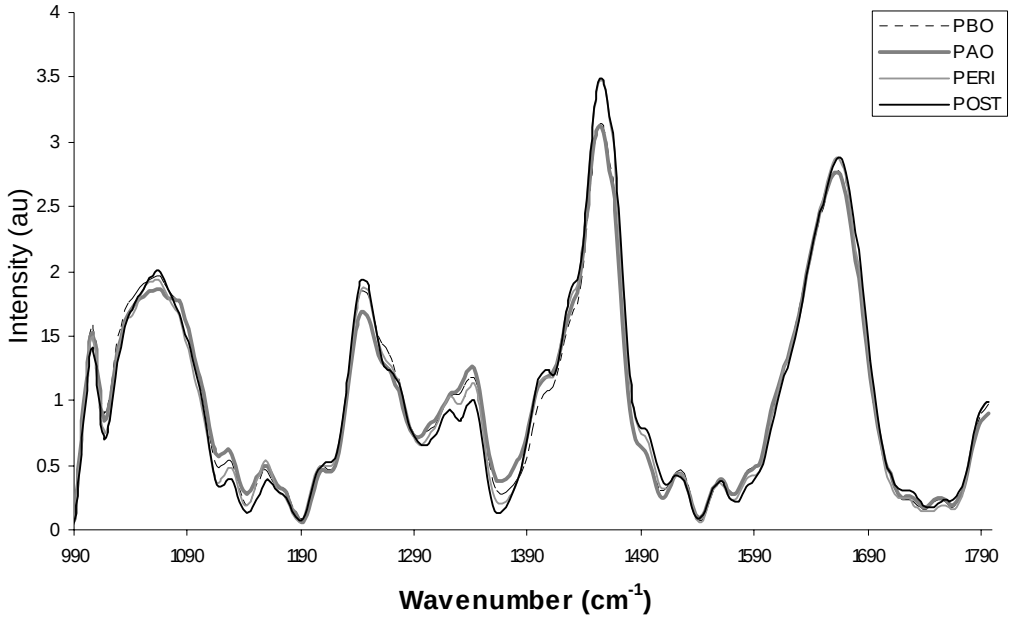


Figure 4.2: Average Raman spectra for post menopausal normal cervix (POST-30), peri menopausal normal cervix (PERI-34), pre-menopausal after ovulation normal cervix (PAO-54) and pre-menopausal before ovulation normal cervix (PBO-47).

Women who were on traditional oral contraceptives were placed in the category in which they best fit. The spectra have subtle but consistent differences, especially at 1250cm^{-1} and 1300-1320cm^{-1}. These differences could be due to the changes caused different levels of estrogen and progesterone during a women's lifetime. As seen in the confusion matrix shown as Table 4.1, the data were classified with an overall accuracy of 98.2%.

Table 4.1: Confusion matrix from the menopausal status data.

		Histopathology			
		PBO	PAO	PERI	POST
Raman Algorithm	PBO	47	0	0	0
	PAO	0	53	0	0
	PERI	0	0	33	1
	POST	0	1	1	28

Only 3 spectra were classified incorrectly, and one of the misclassifications was a post-menopausal spectrum that was atrophic. This is relatively common in post-menopausal women and is clinically treated as normal, although an atrophic cervix has different biochemical components.

4.3.2. Variations in Number of Vaginal Deliveries

The cervix changes significantly during pregnancy, during a vaginal delivery, the cervix must stretch to allow for the baby to leave the uterus. All women who have had a caesarian section (C-section) or a miscarriage were excluded from this study. When comparing the no vaginal deliveries (zero) to one vaginal delivery (one) and two or more (two plus), the classification accuracy was 62%. The confusion matrix is shown in table 4.2.

Table 4.2: Confusion matrix for the number of vaginal deliveries.

		Histopathology		
		Zero	One	Two
Raman Algorithm	Zero	33	1	22
	One	1	13	4
	Two	18	4	36

Due to low classification accuracy, we determined that previous vaginal deliveries do not significantly affect the biochemical composition, as seen in the Raman spectra, of the cervix in the long term and therefore do not need to be included in any of our algorithms. Although variations that would be seen in the Raman spectra still may be present.

4.2.3 Variations in Cervical Health

Variations in previous cervical health have the potential to affect the Raman signatures. Spectra from women who have had previous cervical disease but currently have a healthy cervix (pervious abnormal) and spectra from women who have had no history of cervical disease are shown in Figure 4.3. There are consistent differences at the 1250 cm^{-1} peak and at the shoulder around 1400 cm^{-1}.

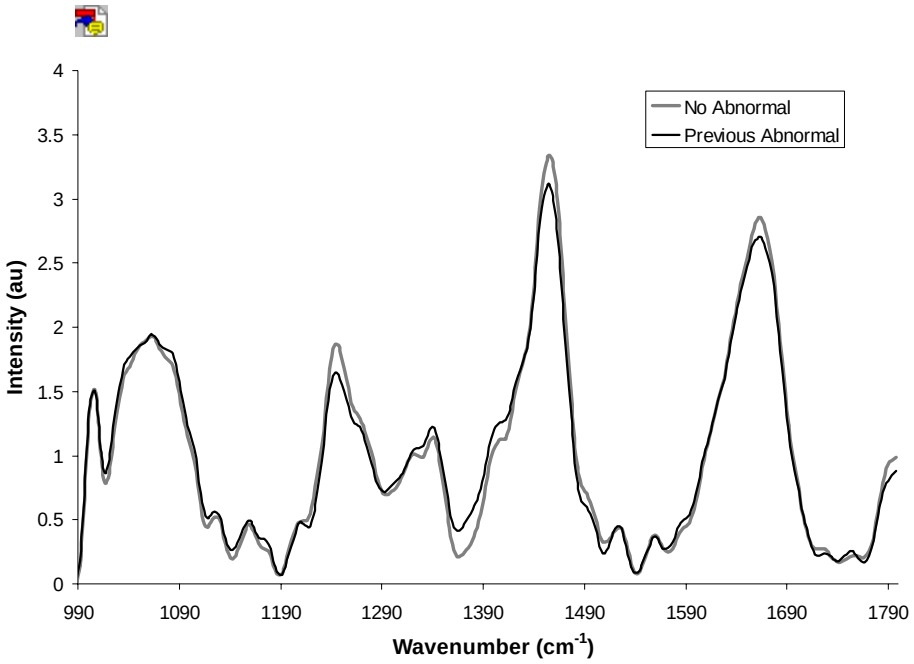


Figure 4.3: Average Raman spectra form no previous abnormal Pap smear (110 spectra and previous abnormal Pap smear (53 spectra)

In table 4.3, the confusion matrix for classification using MRDF and SMLR is shown.

The spectra were classified correctly 99.3% of the time.

Table 4.3: Confusion matrix for previous disease vs. the true normal (no abnormal Pap smear).

		Histopathology	
		True normal	Previous disease
Raman Algorithm	True normal	109	0
	Previous disease	1	53

Although pervious disease has a significant affect on the biochemical signatures of the cervix, this information is not incorporated into the algorithm because no acetic acid has been added to the cervix in these studies. In the data included in the LGSIL to Normal almost all the normal data has had disease somewhere in the cervix, so this difference should not affect the result of the classification algorithm. This information needs to be considered if this technology is used as a screening method but as a guidance of biopsy it does need to be considered because these cervix all have evidence of disease.

4.2.4 Dysplasia Study

Average LGSIL spectra and the average normal spectra are displayed in Figure 4.4. The largest difference occurs between 1230cm^{-1} to 1300cm^{-1}. Many biochemical changes occur as tissue changes from normal cervical tissue to dysplastic cervical tissue, but in LGSIL, these changes are very subtle because only a few of the cells have undergone this transformation. The classification accuracy when discriminating between LGSIL and normal tissue is 81% using MRDF and SMLR (CHAPTER 3).

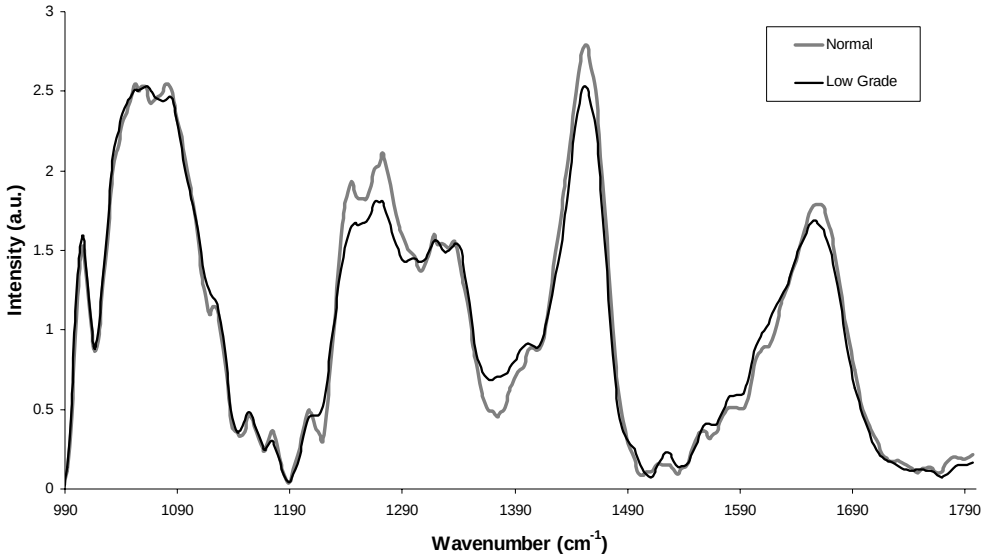


Figure 4.4: Average Raman spectra for normal cervix (34) and low grade cervix (30).

Using the new algorithm that incorporates hormonal status, the classification accuracy improved to 97% for LGSIL. The new higher class discrimination algorithm was developed using the information obtained from the menopausal status study. PERI and POST spectra were not included in this analysis because there were not enough spectra. In table 4.4, the classification using 4 classes: normal-PBO, normal-PAO, low grade-PBO, and low grade-PAO is shown. In this table, there are a total of 4 spectra that are classified incorrectly: two LGSIL and two normal.

Table 4.4: Confusion matrix for LGSIL vs. Normal separated by location in the menstrual cycle.

		Histopathology			
		Normal-PBO	Normal-PAO	LGSIL-PBO	LGSIL-PAO
Raman Algorithm	Normal- PBO	19	0	1	1
	Normal-PAO	0	13	0	0
	LGSIL-PBO	1	0	14	0
	LGSIL-PAO	1	0	0	14

4.3 Comment

This study brings Raman spectroscopy one step closer to clinical use by improving the specificity in diagnosing dysplasia. This improvement was accomplished by incorporating variations in the normal cervix to differentiate LGSIL from normal. It was found that changes due to menopausal state and menstrual cycle location in the normal cervix can be detected with Raman spectroscopy. These stratifications need to be considered when using a classification algorithm to differentiate between normal and dysplastic tissue. When this information is incorporated into the algorithm, the classification accuracy improves from 81% to 97%, indicating the potential of Raman spectroscopy to diagnosis low grade dysplasia. The results are very promising, but the application of this new multi-class algorithm was on a small set of dysplasia/normal data. Both menopausal status as well as location in the menstrual cycle affects the Raman spectra, but due to the nature of cervical dysplasia, only pre-menopausal women were

included in the dysplasia portion of the study. Other studies suggest Raman can be used to diagnose HGSIL with high success [10].

One limitation of this study is that there are some instances where women do not fit into any of the categories (PBO, PAO, PERI or POST). This can be due to a form of birth control (Depo-Provera (Depo)) or to very irregular periods (amenorrhea). The MRDF and SMLR results from the spectral data from patients on Depo were very consistent and are shown in table 4.5. The cervices of the women who use Depo for either birth control or health reasons classifies as pre-menopausal after ovulation (PAO). Since depo delivers a high level of progesterone, it stops the ovaries from releasing eggs, causes the cervical mucus to thicken, and changes the uterine lining, similar to the changes after ovulation. After ovulation, the corpus luteum produces high levels of progesterone, and this progesterone thickens the mucus in the cervix and acts on the lining of the uterus. The amenorrhea results are also as expected. These women have very inconsistent periods and therefore may fall at any point in the cycle. Despite this limitation, this application of Raman spectroscopy would still be very beneficial in the clinic.

Table 4.5: Classification of amenorrhea and depo data as classified by MRDF and SMLR.

		Histopathology	
		Amenorrhea	Depo
Raman Algorithm	PBO	1	0
	PAO	2	4
	PERI	1	0
	POST	2	0

Optical spectroscopy for diagnostics has many advantages, such as real time monitoring, the ability to do "see and treat" procedures, a reduced need for biopsies, and the ability to monitor progression. Although the Raman signatures are significantly weaker than other forms of optical spectroscopy, the slightly longer acquisition time and increased data processing is negated by the increased sensitivity and specificity. For example, using fluorescence spectroscopy, the prospective sensitivity and specificity of a paired multivariate algorithm for discriminating precancers from non-precancerous tissues is 68% [11], whereas Raman spectroscopy has been shown to discriminate these tissues with an accuracy of 97% [12].

Studies show that Raman spectroscopy has the potential to detect small variations both in the normal cervix and in the cervix as it becomes dysplastic [10, 12]. The ectocervix consists of a dense fibromuscular stroma which is composed primarily of collagenous connective tissue and a ground substance of mucopolysaccharides. The connective tissue is approximately 15% smooth muscle and a small amount of elastic

tissue. Hormonal changes, such as menopausal status and location in the menstrual cycle, change the composition of the ectocervix [13]. Therefore, Raman signatures vary significantly depending on location within the menstrual cycle and with the onset and completion of menopause. These differences are important when trying to correctly classify the results. Spectral differences are shown in figure 1; the most notable differences occur around 1250cm^{-1}, 1300cm^{-1}, and 1320cm^{-1}, most likely due to changes in protein levels, especially elastin and collagen. During the menstrual cycle, the cervix becomes softer or more elastic as the level of estrogen increases. After ovulation, this process is reversed, and the cervix loses some of its elasticity. During peri-menopause, the layer of epithelial cells thins, and the vascularity and content of the cervix is erratic, but the spectra are consistent. The most variable and therefore the hardest group to classify is the post-menopausal group. The absence of ovarian estrogen and progesterone causes the cervix to change, including both dryness and atrophy. These conditions are considered normal in a woman who has gone through menopause but could cause major differences in the Raman spectra.

The changes due to dysplasia are different than those associated with hormonal variations. The largest change when comparing LGSIL Raman spectra to normal Raman spectra are in the 1230-1300 cm^{-1} range. This peak range is usually associated with proteins and lipids[14]. It is expected that there will be variations in the protein and lipid content when dysplastic changes occur because of an increase in metabolic activity. Another expected change is a reduction in the glycogen peaks[15] that occur around 1300 cm^{-1}. This difference is expected to be minimal in LGSIL because the disease only

affects a small portion of the epithelium. As the disease progresses towards HGSIL, this drop in the glycogen peak is expected to be more drastic.

Stratification and classification with a non-linear multi-class algorithm yields a posterior probability. This is a powerful tool because it provides the confidence that a particular area is classified correctly. If the confidence is low and the area is suspicious based on the doctor's observation, a biopsy could be taken as current clinical protocol suggests. This method would greatly reduce the number of biopsies taken but would not eliminate them completely. But then, most diagnoses could be determined in the clinic on the day of the visit instead of several days to a week later. This eliminates the stress a woman feels while waiting for results and the need for a follow up visit to discuss the results of the biopsy.

In future work, independent validation on a larger dysplasia patient population will be done to confirm clinical effectiveness. Additionally, other changes in the normal cervix will be characterized and the algorithm modified as needed. Overall, this method for LGSIL diagnosis seems promising and would greatly benefit rural communities and working individuals by limiting the time spent at the doctor's office.

4.4 References

1. American Cancer Society, "Cervical Cancer Resource Center", 2008.

2. G. H. Anderson, D. A. Boyes, J. L. Benedet, J. C. Leriche, J. P. Matisic, K. C. Suen, A. J. Worth, A. Millner and O. M. Bennett, "Organization and Results of the Cervical Cytology Screening-Program in British-Columbia, 1955-85," *British Medical Journal*, 296(6627), 975-978, 1988.

3. M. T. Fahey, L. Irwig and P. Macaskill, "Metaanalysis of Pap Test Accuracy," *American Journal of Epidemiology*, 141(7), 680-689, 1995.

4. K. Nanda, D. C. McCrory, E. R. Myers, L. A. Bastian, V. Hasselblad, J. D. Hickey and D. B. Matchar, "Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: A systematic review," *Annals of Internal Medicine*, 132(10), 810-819, 2000.

5. G. P. Parham, V. V. Sahasrabuddhe, M. H. Mwanahamuntu, B. E. Shepherd, M. L. Hicks, E. M. Stringer and S. H. Vermund, "Prevalence and predictors of squamous intraepithelial lesions of the cervix in HIV-infected women in Lusaka, Zambia," *Gynecologic Oncology*, 103(3), 1017-1022, 2006.

6. E. M. Gill, A. Malpica, R. E. Alford, A. R. Nath, M. Follen, R. R. Richards-Kortum and N. Ramanujam, "Relationship between collagen autofluorescence of the human cervix and menopausal status," *Photochemistry and Photobiology*, 77(6), 653-658, 2003.

7. D. D. Cox, S. K. Chang, M. Y. Dawood, G. Staerkel, U. Utzinger, R. R. Richards-Kortum and M. Follen, "Detecting the signal of the menstrual cycle in fluorescence spectroscopy of the cervix," *Applied Spectroscopy*, 57(1), 67-72, 2003.

8. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Applied Spectroscopy*, 57(11), 1363-1367, 2003.

9. S. K. Majumder, S. Gebhart, M. D. Johnson, R. Thompson, W. C. Lin and A. Mahadevan-Jansen, "A probability-based spectroscopic diagnostic algorithm for simultaneous discrimination of brain tumor and tumor margins from normal brain tissue," *Appl Spectrosc*, 61(5), 548-57, 2007.

10. A. Robichaux-Viehoever, E. Kanter, H. Shappell, D. Billheimer, H. Jones and A. Mahadevan-Jansen, "Characterization of Raman spectra measured in vivo for the detection of cervical dysplasia," *Applied Spectroscopy*, 61(9), 986-993, 2007.

11. N. Ramanujam, M. F. Mitchell, A. MahadevanJansen, S. L. Thomsen, G. Staerkel, A. Malpica, T. Wright, N. Atkinson and R. RichardsKortum, "Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths," *Photochemistry and Photobiology*, 64(4), 720-735, 1996.

12. S. K. Majumder, E. Kanter, A. R. Viehoever, H. Jones and A. Mahadevan-Jansen, "Near-infrared Raman spectroscopy for in-vivo diagnosis of cervical dysplasia: a probability-based multi-class diagnostic algorithm", *Advanced Biomedical and Clinical Diagnostic Systems V,* SPIE, 6430(64300Q-11), 2007.

13. V. L. Katz, R. A. Lobo, G. Lentz and D. Gershenson, "Katz: Comprehensive Gynecology, 5th ed.," 2008.

14. R. A. Bitar, S. Martinho Hda, C. J. Tierra-Criollo, L. N. Zambelli Ramalho, M. M. Netto and A. A. Martin, "Biochemical analysis of human breast tissues using Fourier-transform Raman spectroscopy," *J Biomed Opt*, 11(5), 054001, 2006.

15. J. W. Sellors and R. Sankaranarayanan, *Colpsocopy and Treatment of Cervical Intreaepithelial Neoplasia: A Beginners' Manual*, International Agency for Research on Cancer, Lyon, 2003.

CHAPER V

RAMAN SPECTROSCOPY FOR CERVICAL PRECANCER DETECTION

6.1 Introduction

Cervical cancer is the second most common malignancy in women with over 490,000 cases diagnosed and 274,000 deaths each year [1]. In the last 50 years, early detection has played a role in reducing mortality, but the incidence of pre-invasive cervical squamous carcinoma has risen dramatically [2]. While annual Pap smears are the standard of care amongst women in developed countries, most women in Asia and Africa typically do not have one. Due to the lack of consistent screening techniques, the percentage of women that die from invasive cervical cancer increases from 32% in the US to 56% in the entire world. Furthermore, over 80% of cervical cancer cases occur in developing nations [3]. Invasive cervical cancer is the number one cause of cancer related deaths among women in sub-Saharan Africa [3]. The Pap smear is not the standard of care in developing countries because of high costs, difficulty in preserving cell samples and transporting slides, lack of trained lab technicians to analyze the slides, and obstacles in bringing the women back for follow-up tests and for treatment [4]. Thus a tool that can screen and diagnose cervical cancer would relieve a tremendous and unnecessary burden in developing countries. A tool that has the ability to accurately diagnose cervical cancer must be sensitive enough to differentiate between various stages, including normal, inflammation, metaplasia, low grade squamous intraepithelial lesion (LGSIL), high grade squamous intraepithelial lesion (HGSIL), and cancer. From

this diagnosis, the care provider could decide on the appropriate treatment, whether to do nothing, to perform a loop electrosurgical excision procedure (LEEP), or something more drastic like a hysterectomy.

Optical techniques have the potential to fill this need. These techniques can be used as a "see and treat" method for detecting abnormalities in the cervix due to their noninvasive nature and ability to detect both biochemical (Raman [5-9] and fluorescence spectroscopy [10, 11]) and structural (optical coherence tomography [12]) changes in the cervix. Raman spectroscopy is a molecular specific, noninvasive technique that measures the biochemical composition of a molecule by inducing vibrational or rotational transitions [13]. It has been used for many years to probe into the biochemistry of various biological molecules [8]. Recently, there has been an interest in using this technique for diagnosing precancers and cancers [14].

Although only a limited number of biological molecules contribute to tissue fluorescence, several biological molecules such as nucleic acids, proteins, and lipids have distinctive Raman features that yield structural and environmental information. Results indicate that molecular and cellular changes that occur in precancerous tissues as well as in benign abnormalities such as inflammation yield characteristic Raman features that allow their differentiation. For example, one of the more prominent changes that occur with cancer and precancer is increased cellular nucleic acid content; extensive DNA studies indicate that it may be possible to sample this change using Raman spectroscopy [15]. On the basis of these biochemical differences, several groups have studied the potential of vibrational spectroscopy for cancer diagnosis in various organ sites [8]. These groups have shown that features of the vibrational spectra can be related to

molecular and structural changes associated with neoplastic transformation. Raman spectroscopy has been applied towards *in vitro* detection of epithelial and mesenchymal cancers such as breast, colon, esophagus and gynecologic tissues [14]. Recent developments in detector and source technologies have resulted in acquiring Raman spectra from tissue in 1-3 seconds. Several fiber optics probes have been developed that are capable of measuring Raman spectra *in vivo* making it possible to apply this technique in a clinical setting. An increased number of reports have been published on applying Raman spectroscopy to detect cancers in organs *in vivo* such as the cervix, skin, breast and gastrointestinal tract with high sensitivities and specificities [9, 16-19]. In developing countries such as India and Zambia, where the use of the Pap smear is not practical for cervical cancer detection, Raman spectroscopy can be utilized as a screening and diagnostic tool.

Raman spectroscopy does have a couple drawbacks, but these can easily be overcome. First, measurements must be taken in the dark as sunlight affects the spectra. To address this concern, the procedure can be done in a room without windows, lit with incandescent lights. Secondly, the signal output tends to be weak. Post processing is done to extract weak Raman signal from the much stronger fluorescence.

Raman measurements can be used for both screening and diagnosis. In the US, women are usually screened every year with the Pap smear. An abnormal diagnosis from the Pap smear is usually followed by HPV testing and/or a colposcopy guided biopsy. Our Raman system can be used in both settings, replacing Pap smear and biopsy. In this paper, we show a successful clinical application of Raman spectroscopy to diagnose and screen for cervical cancer and precancer.

5.2 Methods

5.2.1 Clinical Study Design-Dysplasia Patients

Forty-three patients undergoing a colposcopy guided biopsy were recruited to participate in the study as approved by the Institutional Review Board (IRB). To be eligible for enrollment, the patient must be undergoing a colposcopy guided biopsy, be between the ages of 18-75 and still have a cervix (no history of a hysterectomy). After informed consent was obtained from each patient, the cervix was exposed and visually examined by the attending physician. Acetic acid was applied to the cervix (to turn abnormal areas white for visualization) and abnormal tissue was removed and placed in fixative solution for pathology examination. After applying acetic acid and before removing the tissue, spectra were acquired from multiple areas of tissue to be removed and 1-2 visually normal areas. Based on the pathology, spectra were placed into three categories for analysis: no disease (normal), LGSIL, and HGSIL. Only pre-menopausal patients were included in the analysis.

5.2.2 Clinical Study Design-Pap Smear Patients

Twenty-nine patients undergoing a routine pap smear were recruited to participate in the study as approved by the IRB. To be eligible for enrollment, the patient must be undergoing a routine pap smear, be between the ages of 18-75 and still have a cervix (no history of a hysterectomy). Informed consent was obtained from each patient prior to the procedure. The cervix was exposed and visually examined by the attending physician. The pap procedure was done according to standard clinical protocol. The cervix was

wiped clean with a dry cotton swab and then with a saline solution. Acetic acid is applied to the cervix and measurements were taken in the same location as before. The spectra were considered normal if the Pap smear was negative. The patient's age, last period, artificial hormones, any pervious abnormal Pap smears and menopausal status were all noted upon chart review. Only pre-menopausal patients with no previous abnormal Pap smears were included in this study. These patients are referred to as true normal. Table 5.1 lists the four categories used to describe the data set.

Table 5.1: Summary of the categories used to describe this data set.

Description	True Normal	Normal	LGSIL	HGSIL
History of Abnormal Pap smear (i.e. one or more)	No	Yes	Yes	Yes
Evidence of Disease on the cervix regardless of location	No	Yes	Yes	Yes
Evidence of disease where measurement is taken	No	No	Yes	Yes
Presence of acetic acid	Yes	Yes	Yes	Yes
Biopsy results	NA	No evidence of disease	CIN I, HPV cellular effects	CIN II, CIN III, CIS

5.2.3 Data Collection

Raman spectra were collected from multiple sites *in vivo* using a portable Raman spectroscopy system consisting of a 785 nm diode laser (Process Instruments, Inc., Salt Lake City, UT), beam-steered fiber optic probe (Visionex Inc., Atlanta, GA), imaging spectrograph (Kaiser Optical Systems, Inc., Ann Arbor, MI), and back-illuminated, deep-depletion, thermo electrically cooled charge coupled device (CCD) camera (Roper Scientific, Inc., Princeton, NJ), all controlled with a laptop computer. Details of the system have been previously reported [20]. The fiber optic probe delivered 80mW of incident light onto the tissue for 3 seconds. For each measurement, overhead fluorescent lights were turned off.

Spectral calibration of the system was preformed each day using a neon-argon lamp, naphthalene and acetaminophen standards to correct for system wavenumber, laser excitation, and throughput variations. The spectra were processed for fluorescence subtraction and noise smoothing using the modified mean method, described previously [20]. Following data processing, each spectrum was normalized to its mean spectral intensity across all Raman bands to account for overall intensity variability. These normalized spectra were categorized according to menopausal status and histology as determined by the pathologist.

5.2.4 Statistical Analysis

The analysis technique used in this paper has been described elsewhere [21]. The process consists of two steps - first, extraction of diagnostic features from the spectra using the nonlinear maximum representation and discrimination feature (MRDF); second,

developing a probabilistic scheme of classification based on linear sparse multinomial logistic regression (SMLR) for classifying nonlinear features into corresponding tissue categories. All classification was done using leave-one-patient out cross-validation.

5.3 Results and Discussion

Raman spectroscopy has been shown to be an effective method for cervical dysplasia detection. Previous studies show that Raman can be used to differentiate between high grade dysplasia and normal tissue. Due to inadequate low grade data, the true ability of this method had not been shown. This study provides evidence that Raman spectroscopy has the potential to differentiate between LGSIL, HGSIL and normal cervix.

The average Raman spectra from true normal ectocervix, normal ectocervix, LGSIL and HGSIL are shown in Figure 5.1. The biggest difference can be seen between dysplasia (both high grade and low grade) and normal in the 1200-1300 cm^{-1} range, highlighted with a dashed box. The peak around 1250 cm^{-1}, usually associated with collagen, is higher in both true normal and normal spectra. Conversely, the peak around 1330 cm^{-1} is higher in the LGSIL spectrum and in the HGSIL spectrum; this peak is usually associated with DNA and glycogen [22, 23]. This difference is expected because as tissue gets more dysplastic, the amount of cellular DNA will increase due to rapid dividing and irregular growth [15].

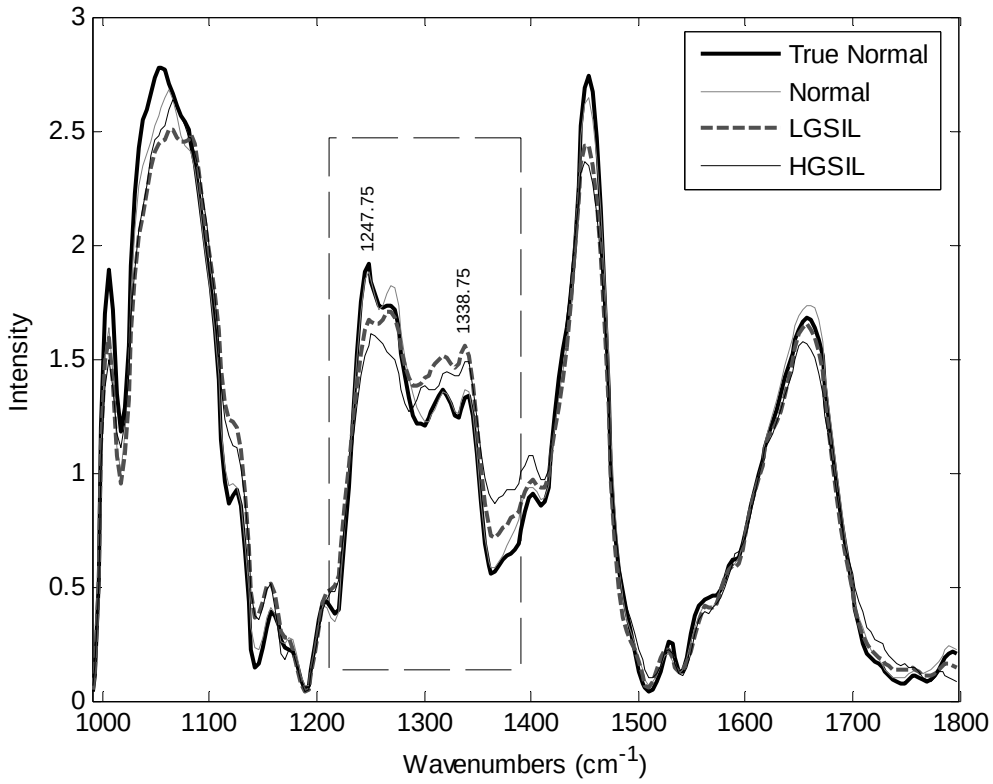


Figure 5.1: Average Raman spectra for true normal ectocervix, normal ectocervix, LGSIL and HGSIL.

A four class algorithm (MRDF and SMLR) was used to classify the data (true normal, normal, LGSIL and HGSIL). The classification results were obtained based on leave-one patient out cross validation of the entire data set. This algorithm was able to classify dysplastic cervix with 97% accuracy. Its best performance was with true normal and HGSIL where all the spectra were classified correctly, and it misclassified approximately 2-9% in the other 2 categories. The result of the classification algorithm can be seen in Table 5.2.

Table 5.2: Classification of samples using on MRDF and SMLR leave one patient out cross validation.

		Histological Classification			
		True Normal	Normal	LGSIL	HGSIL
Raman Classification	True Normal	52	1	1	0
	Normal	0	35	1	0
	LGSIL	0	0	43	0
	HGSIL	0	1	0	12

A total of 4 spectra classified incorrectly suggesting several limitations of this current method. Two normal samples classified incorrectly: one as true normal and one as HGSIL; two LGSIL samples classified incorrectly: one as true normal and one as normal. For example, the biopsy correlating to one of the normal spectra that classified incorrectly came back as intense chronic endocervicitis suggesting that not all benign cervical conditions will classify as normal and that inflammation should be a separate category in the algorithm. The other normal spectra that classified incorrectly classified as true normal was taken from a patient who had never had an abnormal Pap smear, the area that was misclassified was believed to be normal. For the LGSIL sample that classified as true normal, biopsy results suggested mild HPV changes (LGSIL) in the tissue but no CIN, for this particular patient, this was the 1st abnormal Pap smear and the patient was young. These two misclassifications suggest that our method may not be able

to detect initial local HPV effects or the initial diffuse HPV effects. The other LGSIL that misclassified as normal had mild HPV changes where the biopsy was taken but did have CINI elsewhere, it is suspected that the biopsy was not taken exactly where the measurement was taken and could have possibly been normal. In figure 5.2, the posterior probabilities are shown from the data in Table 5.1. This displays how likely a particular spectrum will classify into each of the categories, which may be useful clinically because a doctor can determine the confidence that a measurement fits within a certain category.

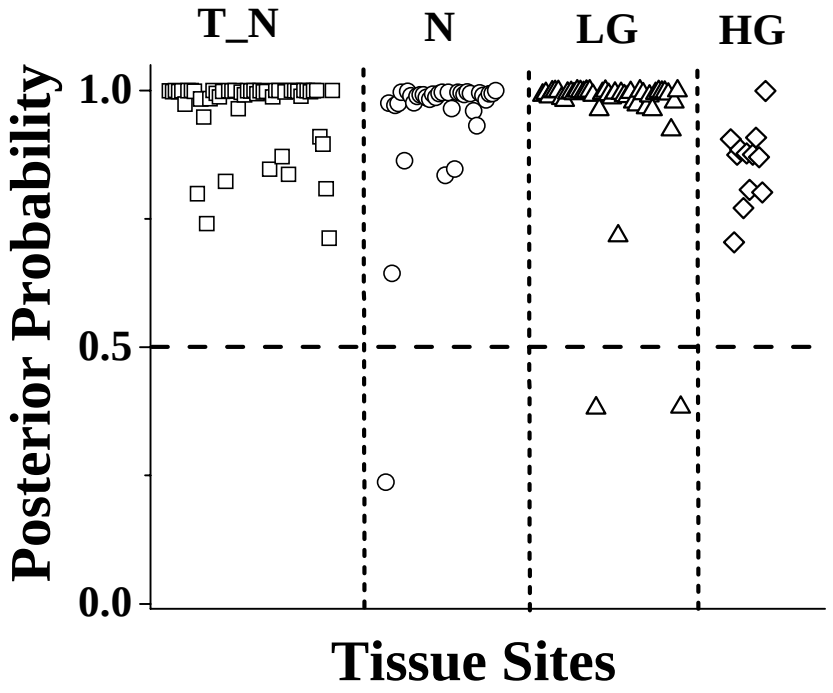


Figure 5.2: Posterior probabilities of classification as true normal ectocervix (T_N), normal ectocervix (N), LGSIL (LG) and HGSIL (HG).

Since menopausal stratus affects Raman spectra, only pre-menopausal women were considered for this study (unpublished work). It is expected that if the algorithm

was run on a group of post or peri-menopausal women, we would find similar results. As most cases of cervical dysplasia occur in younger women, our data set from older women is not large enough at this time to carry statistical significance.

Although spectral differences are small and cannot easily be seen, they are significant enough to be picked up by sophisticated algorithms. There are many advantages to using MRDF and SMLR for data reduction and classification. One advantage is that it outputs a posterior probability as a measure of confidence in correctly classifying the tissue. Higher posterior probabilities lead to more confidence in the result, allowing the doctor to be more confident in their diagnosis. Although the primary application of this system is to "see and treat" in developing countries, this method could also be used in junction with colposcopy to reduce the number of biopsies taken. For example, a doctor could place the probe on the suspected area and find a posterior probability of 97% that the tissue is LGSIL. They may consider this area low grade without having to biopsy. In situations where the area has a lower posterior probability of being low grade, the doctor may take a biopsy to ensure that the area is low grade.

This method would be very useful in developing countries, where "see and treat" methods are optimal. One major problem with screening alone is poor follow-up testing among women with abnormal pap smears. Usually, an abnormal Pap smear requires a follow-up biopsy and then a return visit 3-6 months later depending on the result. However, an estimated 10-61% of women with abnormal Pap smears do not show up for follow-up testing [24]. Factors associated with this noncompliance include (1) only an elementary education, (2) prior surgery, (3) additional diseases, (4) consumption of medications for chronic conditions, and (5) family illness [25]. Additionally, only an

estimated 5 percent of women in developing countries have been screened for cervical dysplasia in the past 5 years compared with some 40 -50 percent of women in developed countries [25]. In developing countries, standard practice is for a nurse to photograph the cervix to send to a doctor for diagnosis. This time-consuming process also has a high error rate and level of ambiguity. Our method would allow a nurse to find suspicious areas, take Raman measurements, report an accurate diagnosis, and decide on treatment all on the same day, reducing the number of patients who are treated unnecessarily and ensuring patients receive an accurate diagnosis the same day in the clinic. Raman spectroscopy has the potential to solve many of the obstacles facing accurate diagnosis and screening of cervical cancer.

5.5 Acknowledgements

The authors would like to acknowledge the financial support of the NCI/NIH (R01-CA95405). We would also like to thank the doctors and staff at Vanderbilt University Medical Center and Tri-state Women's Health for all their help.

5.6 References

1. G. P. Parham, V. V. Sahasrabuddhe, M. H. Mwanahamuntu, B. E. Shepherd, M. L. Hicks, E. M. Stringer and S. H. Vermund, "Prevalence and predictors of squamous intraepithelial lesions of the cervix in HIV-infected women in Lusaka, Zambia," *Gynecologic Oncology*, 103(3), 1017-1022, 2006.

2. L. Burke, D. A. Antonioli and B. S. Ducatman, *Colposcopy, text and atlas*, Appleton and Large, Norwalk, 1991.

3. J. H. Farley, J. F. Hines, R. R. Taylor, J. W. Carlson, M. F. Parker, E. R. Kost, S. J. Rogers, T. A. Harrison, C. I. Macri and G. P. Parham, "Equal care ensures equal survival for African-American women with cervical carcinoma," *Cancer*, 91(4), 869-73, 2001.

4. "Noncommunicable Disease is South-East Asia Region: A profile," 2002.

5. A. Viehoever Robichaux, E. Kanter, H. Shappell, D. Billheimer, H. Jones and A. Mahadevan-Jansen, "Characterization of Raman Spectra Measured In vivo for the Detection of Cervical Dysplasia," *Applied Spectroscopy*, 61(9), 986-993, 2007.

6. A. Robichaux, H. Shappell, B. Huff, H. Jones and A. Mahadevan-Jansen, "In vivo detection of cervical dysplasia using near infrared Raman spectroscopy," *Lasers in Surgery and Medicine*, 3-3, 2002.

7. S. K. Majumder, E. Kanter, A. R. Viehoever, H. Jones and A. Mahadevan-Jansen, "Near-infrared Raman spectroscopy for in-vivo diagnosis of cervical dysplasia: a probability-based multi-class diagnostic algorithm", *Advanced Biomedical and Clinical Diagnostic Systems V*, SPIE, 6430(64300Q-11), 2007.

8. A. Mahadevan-Jansen and R. Richards-Kortum, "Raman Spectroscopy for the detection of cancers and precancers," *J Biomed Opt*, 1(1), 31-70, 1996.

9. A. Mahadevan-Jansen, W. F. Mitchell, N. Ramanujam, U. Utzinger and R. Richards-Kortum, "Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue in vivo," *Photochemistry and Photobiology*, 68(3), 427-431, 1998.

10. B. M. Pikkula, O. Shuhatovich, R. L. Price, D. M. Serachitopol, M. Follen, N. McKinnon, C. MacAulay, R. Richards-Kortum, J. S. Lee and E. N. Atkinson, "Instrumentation as a source of variability in the application of fluorescence spectroscopic devices for detecting cervical neoplasia," *Journal of Biomedical Optics*, 12(3), 2007.

11. J. A. Freeberg, J. L. Benedet, C. MacAulay, L. A. West and M. Follen, "The performance of fluorescence and reflectance spectroscopy for the in vivo diagnosis of cervical neoplasia; point probe versus multispectral approaches," *Gynecologic Oncology*, 107(1), S248-S255, 2007.

12. P. F. Escobar, L. Rojas-Espaillat, S. Tisci, C. Enerson, J. Brainard, J. Smith, N. J. Tresser, F. I. Feldchtein, L. B. Rojas and J. L. Belinson, "Optical coherence tomography as a diagnostic aid to visual inspection and colposcopy for preinvasive and invasive cancer of the uterine cervix," *International Journal of Gynecological Cancer*, 16(5), 1815-1822, 2006.

13. R. J. Colthrup, *Infrared and Raman spectroscopy*, 1991.

14. E. B. Hanlon, R. Manoharan, T. W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari and M. S. Feld, "Prospects for in vivo Raman spectroscopy," *Physics in Medicine and Biology*, 45(2), R1-R59, 2000.

15. M. S. Feld, R. Manoharan, J. Salenius, J. Orenstein-Carndona, T. J. Roemer, J. F. Brennan III, R. R. Dasari and Y. Wang, "Detection and characterization of human tissue

lesions with near-infrared Raman spectroscopy", *Advances in Fluorescence Sensing Technology II*, SPIE, 2388(99-104), 1995.
16. M. V. P. Chowdary, K. K. Kumar, J. Kurien, S. Mathew and C. M. Krishna, "Discrimination of normal, benign, and malignant breast tissues by Raman spectroscopy," *Biopolymers*, 83(5), 556-569, 2006.

17. T. R. Hata, T. A. Scholz, I. V. Ermakov, R. W. McClane, F. Khachik, W. Gellermann and L. K. Pershing, "Non-invasive Raman spectroscopic detection of carotenoids in human skin," *Journal of Investigative Dermatology*, 115(3), 441-448, 2000.

18. M. G. Shim, L. Song, N. E. Marcon and B. C. Wilson, "In vivo near-infrared Raman spectroscopy: Demonstration of feasibility during clinical gastrointestinal endoscopy," *Photochemistry and Photobiology*, 72(1), 146-150, 2000.

19. A. P. Oliveira, R. A. Bitar, L. Silveira, R. A. Zangaro and A. A. Martin, "Near-infrared Raman spectroscopy for oral carcinoma diagnosis," *Photomedicine and Laser Surgery*, 24(3), 348-353, 2006.

20. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Applied Spectroscopy*, 57(11), 1363-1367, 2003.

21. S. K. Majumder, S. Gebhart, M. D. Johnson, R. Thompson, W. C. Lin and A. Mahadevan-Jansen, "A probability-based spectroscopic diagnostic algorithm for simultaneous discrimination of brain tumor and tumor margins from normal brain tissue," *Appl Spectrosc*, 61(5), 548-57, 2007.

22. C. Kendall, N. Stone, N. Shepherd, K. Geboes, B. Warren, R. Bennett and H. Barr, "Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus," *J Pathol*, 200(5), 602-9, 2003.

23. A. Nijssen, T. C. Bakker Schut, F. Heule, P. J. Caspers, D. P. Hayes, M. H. Neumann and G. J. Puppels, "Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy," *J Invest Dermatol*, 119(1), 64-9, 2002.

24. E. Shinn, K. Basen-Engquist, T. Le, A. Hansis-Diarte, D. Bostic, J. Martinez-Cross, A. Santos and M. Follen, "Distress after an abnormal Pap smear result: scale development and psychometric validation," *Prev Med*, 39(2), 404-12, 2004.

25. J. Bornstein and H. Bahat-Sterensus, "Predictive factors for noncompliance with follow-up among women treated for cervical intraepithelial neoplasia," *Gynecol Obstet Invest*, 58(4), 202-6, 2004.

CHAPER VI

CONCLUSIONS AND FINAL REMARKS

6.1 Summary of Conclusions

Many discoveries were made in the course of this dissertation. In this work, Raman spectroscopy was applied in vivo to over 200 patients. The conclusions made supported by this dissertation are listed below.

1) Multi-class discrimination algorithms can successfully classify cervical dysplasia into four separate categories demonstrating sensitivity and specificity greater than 90%, exceeding both more traditional binary methods (PCA and LDA) and the current gold standard of colposcopy.
2) Changes due to menopausal state and menstrual cycle location in the normal cervix can be detected with Raman spectroscopy. This is due to hormonal changes that change the composition of the ectocervix.
3) Changes due to previous vaginal births are not detectable by Raman spectroscopy but changes due to previous disease are.
4) The cervix of women who use Depo for either birth control or health reasons classifies as pre-menopausal after ovulation (PAO), since Depo delivers a high level of progesterone.

5) When comparing Raman spectra from pre and post menopausal women, there is a need for stratification based on menopausal status for effective diagnosis in patients with evidence of cervical disease.

6) When this information is incorporated into the algorithm, the classification accuracy improves from 81% to 97% indicating the wide-spread potential of Raman spectroscopy.

7) Finally, using true normal, normal of disease, HGSIL and LGSIL, the true capability of Raman spectroscopy was tested, the classification accuracy was 97%. Therefore I believe that Raman spectroscopy has the potential to solve many of the obstacles facing accurate diagnosis and screening of cervical cancer.

This work resulted in 1 accepted peer-reviewed publication, 2 submitted peer-reviewed publications (plus 1 in preparation) and 3 conference presentations/abstracts. This work also resulted in 3 second author publications.

6.2 Contributions to the field and Future directions

Over the course my PhD work, I have demonstrated to many people that there might be a different way to detect cervical abnormalities. I have shown many doctors the possible see and treat application of this method that will be particularly usefully in developing countries such as sub-Sahara Africa and India. Although I was unable to implement this during my PhD work, I did bring the research to a point that this could be done in the near future. I was able to show with high success, that Raman can be used to

detect cervical dysplasia; now the only issues are political and bureaucratic to get this done elsewhere.

This method would be very useful in developing countries, where “see and treat” is an optimal method. One major problem with screening alone is poor follow-up testing among women with abnormal Pap smears. The majority of the time, an abnormal Pap smear requires a follow-up biopsy and then a return visit 3-6 months later depending on the result of the biopsy. However, an estimated 10 -61 percent of women with abnormal Pap smears do not show up for follow-up testing [1]. Factors associated with this noncompliance include (1) only an elementary education, (2) prior surgery, (3) additional diseases, (4) consumption of medications for chronic conditions, and (5) family illness [2]. Additionally, only an estimated 5 percent of women in developing countries have been screened for cervical dysplasia in the past 5 years compared with 40 -50 percent of women in developed countries. One of the current practices in developing countries is for a nurse to take a photograph of the cervix and send it to a doctor to get his diagnosis. This takes time and has many sources of error and ambiguity. Our method would allow a nurse to find a suspicious area, take a Raman measurement of the area, and then effectively diagnosis the area. The nurse could then either allow the women to go home untreated, treat the lesion with a LEEP in the clinic that day or send the patient to a doctor if the diagnosis is more serious. This process could reduce the number of patients who are treated unnecessarily and ensure patients are getting an accurate diagnosis during the same day in the clinic. One major issue in both the lower socioeconomic classes and in undeveloped countries is education. We need to educate women that cervical screening is very important because it has no symptoms in its early stages and it is caused

by a sexually transmitted disease- HPV. HPV is the most common STD and the majority of the population who have HPV are unaware.

There are a few things that need to be done in order to have this technology implemented. First, we need to understand if there are any differences due to race or ethnicity. Most of the data I collected have been from the Caucasian population due to populations at Vanderbilt and in Northern Kentucky. Although I don't expect there to be any differences due to race or ethnicity, this does need to be quantified before the technology can be extrapolated to the developing world. An IRB has been recently approved at Meharry Medical College in Nashville, in order to determine if there any effects on the Raman spectra due to race or ethnicity. If any differences exist in the baseline Raman spectra of these populations, it is possible that they are not due to race but are more due to culture and socioeconomic class leading to nutritional, hygienic, or cultural differences. We may or may not be able to easily quantify these differences (if any exist) but need to consider this aspect when attempting to move this technology across cultures.

Although I was able to significantly reduce the size of the system and make the system more clinically applicable, some changes still need to be made. One thing that needs to be done is the process needs to be automated, so spectra can be taken, processed, and placed directly into the discrimination and algorithm, resulting in a posterior probability and a diagnosis. This would be particularly useful in clinics in developed countries as this method may be used in conjunction with biopsy to offer an additional confidence level. For example, if the output of this algorithm is 60% low grade but we

are suspicious that it could be high grade due to the doctor's colposcopic examination or high grade Pap smear results, a biopsy can still be taken to confirm the results. Unfortunately, there are some difficulties in combining these programs into one that can easily be run. The software will have to rewritten in a format that can be either run through Matlab or Labview. My programming skills are not sophisticated enough to implement this change and hopefully we can find a programmer to make this a reality. Also, additional changes are still being added to the algorithm to make it even more sensitive and specific.

The mapping study was very informative. Many key findings were made. We were able to see how much of the stoma we capture in our probe spectra. Unfortunately, it was very difficult to obtain dysplasia samples. There are only a few dysplasia patients included in the study due very limited availability. Most places are unwilling to part with any portion of the tissue, even just for a few hours because of its diagnostic potential. I think that a longer study (over several years) that allows for more dyspastic tissue samples to be studied and therefore more statistically sound observations could be made. This would allow us to confidently quantify the differences that are observed in our probe based study. If we could say the increase in the 1330 cm^{-1} peak is due to the increased number of nuclei or larger nuclei or intercellular glycogen, we would better understand the biochemical changes that occur as the cervix becomes dysplasic and what we are actually detecting when using Raman spectroscopy.

One thing I found particularly interesting is that it was necessary to train the physician in using the device. Taking spectra on their hands was not adequate to get the patient experience. We found that data did not classify well in either study for each of

the doctors' 1st measurement. In the normal study, slipping was a major issue due to the smaller speculum used as well as the lack of acetic acid to dry out the cervix. There is a characteristic peak that is not in the other spectrum that occurs around 1290 cm-1 when the probe slips during measurement. In the dysplasia study, the differences were not as obvious, but data were consistently misclassified on the physicians 1st day of measurements. This could be due to the probe being held on the tissue at angle. It might be necessary to have a sensor at the end of the probe to ensure it is in full contact. This issue was easily resolved after the doctor completed several measurements.

If I could change one thing, I would have kept my normal study consistent with the dysplasia study. In the normal study, acetic acid was not used was not used for consent purposes; I thought this additional step may limit the number of patients that we were able to consent because of the additional time and slight discomfort. Now, an additional number of patients need to be collected so the true normal can be included in the discrimination algorithm. There are some differences in the cervix after any sort of dysplastic change and this case needs to be considered when developing a diagnostic technique. However, depending on the utilization of this method in the clinic, the addition may not be fully necessary.

No matter who undertakes this project, I think it will greatly benefit society. I believe that Raman spectroscopy will be used in the future to detect cervical dysplasia especially in developing areas. I am excited to see where this projects goes, whether or not I am directly involved or just catching up by what is published in the literature. I hope to be involved at some level especially in the next few years. I cannot wait to see not only the future of Raman spectroscopy in detecting cervical dysplasia, but also how

optical spectroscopy and optical imaging revolutionizes cancer detection and cancer treatment and medicine in general. I feel that I have done my part in advancing optical technologies in medicine. I feel that I have opened eyes to what optical technologies are and what their capabilities may be. We have seen a few of the optical technologies (OCT) move into the medical norm and I feel that some of these other modalities are almost ready to make the jump from the lab to the clinic.

Graduate school has been an enlightening experience; not only allowing me to have a greater understanding of BME but also of myself. In my project, I was able to develop professional relationships. I think getting my PhD will open many doors for me and what I have learned while at Vanderbilt University is irreplaceable. The amount of growth emotionally, mentally, and professionally would not have been possible without my graduate experience at Vanderbilt.

6.3 References

1. E. Shinn, K. Basen-Engquist, T. Le, A. Hansis-Diarte, D. Bostic, J. Martinez-Cross, A. Santos and M. Follen, "Distress after an abnormal Pap smear result: scale development and psychometric validation," *Preventive Medicine*, 39(2), 404-412, 2004.

2. J. Bornstein and H. Bahat-Sterensus, "Predictive factors for noncompliance with follow-up among women treated for cervical intraepithelial neoplasia," *Gynecologic and Obstetric Investigation*, 58(4), 202-206, 2004.

APPENDIX I

INSTUMENTATION VARIATIONS

A1.1 Raman Instrumentation

A Raman spectroscopy system that is used for clinical applications consists typically of 5 major components: (1) a fiber-optic probe, (2) a laser, (3) a spectrograph, (4) a charged-coupled device (CCD), and (5) a computer. A typical configuration (and the one used in our lab) is shown in Figure A1.1.

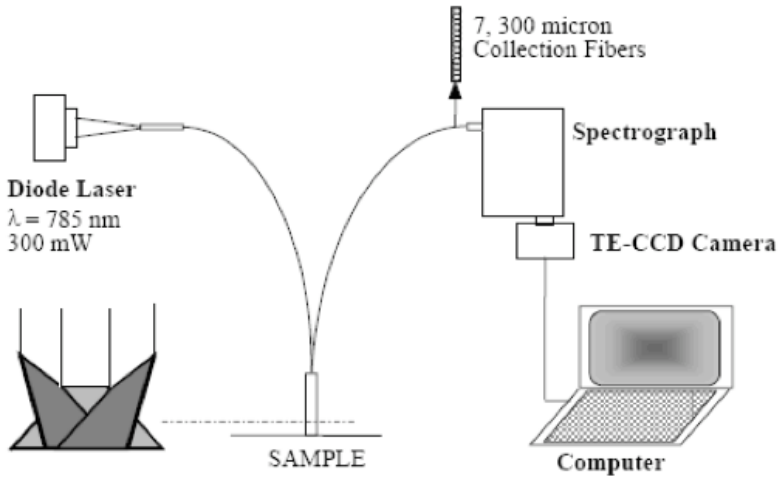


Figure A1.1: Schematic of a typical Raman spectroscopy system used in clinical applications.

The probe is used for both excitation and collection. The probe used in all of the experiments described in this dissertation was built by Visonex (Atlanta, GA); it is made of beam-steered fused silica with a total of 7 collection fibers that surround a single

emission fiber. A diode laser with near-IR excitation wavelengths are used most often in clinical applications because of their size and portability. A 785nm laser made by Process Instruments, Inc. (Salt Lake City, UT) has been used for all the studies, although we have recently found that a smaller diode laser made by Innovative Photonics Systems (Monmouth, NJ) has a comparable signal and is much smaller. The collected light is directed into f-number matched holographic spectrograph in order to provide high throughput and a flat image field at the detector. The spectrograph used to collect the data discussed here is made by Kaiser Optical System, Inc. (Ann Arbor, MI). A CCD is composed of rectangular arrays of photosensitive pixels arranged either in horizontal or vertical directions. For Raman spectroscopy, the wavelength direction corresponds to the horizontal rows, and the columns are binned vertically so that an intensity corresponds to each wavelength. Our system employs a back-illuminated, deep depletion CCD. The current configuration has a thermoelectrically (TE) cooled CCD camera made by Roper Scientific, Inc. (Princeton, NJ), but previous versions of the system have used a liquid nitrogen (LN) cooled camera also made by Roper Scientific. From here on out, these will be referred to as TE and LN. A photograph of each of these systems is shown in figure A1.2, the system with the LN camera is in figure A1.2.a and the one with the TE is in A1.2.b.

(a) (b)

Figure A1.2. Pictures of the systems described in this chapter: (a) the LN system and (b) TE system.

Interchanging the camera throughout clinical measurements has the potential to create spectral variation that lowers the integrity of the measured spectra. But the change was made from the LN to the TE camera to improve clinical applications. The LN camera requires that liquid nitrogen be put into the camera every few days to keep the camera cold. Also, note the size and the appearance differences that came with the change of the cameras, the one in Figure A1.2.b has much more practical clinical application. After the liquid nitrogen was put in, it took over an hour for the camera to be cooled to the appropriated temperature (-70C). The camera also had a controller that had to be attached to a battery pack at all times to maintain the temperature. This made the system more bulky and more difficult to transport into the clinic. The new TE camera is USB controlled and can cool to -70 C in less than 15 minutes. This system does not have to be attached to a battery pack (although ideal when trying to move from room to room in the clinic) and has no controller. This camera variation must be understood and then corrected for in order to increase the diagnostic reliability of Raman spectroscopy.

A1.2 Variations in Instrumentation

In order for Raman spectroscopy to be an effective diagnostic tool, the origins and relative contributions of spectral variation must be understood. The goal of this PhD project is not to quantify these spectral variations, but, because a few became apparent as the project progressed, the differences were noted and changes were made. There have been several groups that have studied the reproducibility of the Raman spectra of healthy tissue, in particular, skin [1] and cervical tissue. [2] This suggests that spectra recorded from the same sample should be similar a posses minimal variability but, as you can see in Figure A1.3, this is not the case when two different Raman spectroscopy instrumentations measured normal spectra of the skin. [3, 4]

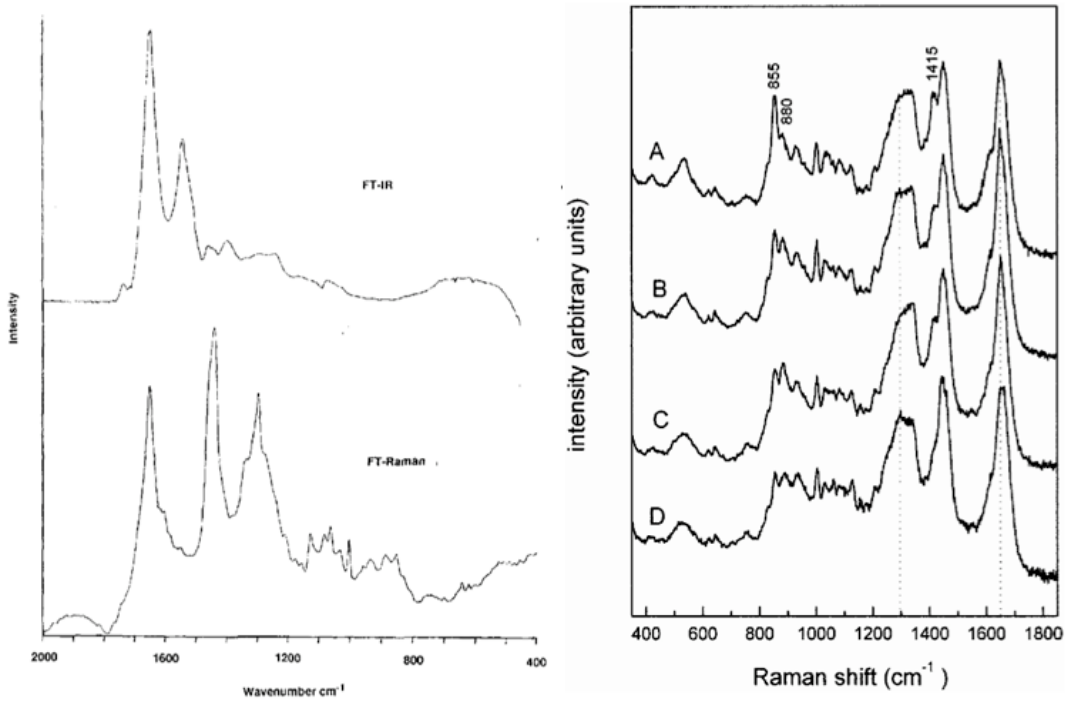


Figure A1.3: Raman skin spectra recorded with different Raman spectroscopy instrumentation systems.

In general, the spectral shape is similar, but there are regions of variation; in this particular case, the signature is slightly different between 800 and 1500 cm^{-1}. Therefore, it is necessary to be conscious of variations in data caused by instrumentation; these variations occur day-to-day as well as over the entire period of a study. Because the differences shown in Figure A1.3 are from different institutions with different instrument configuration, differences in more similar systems may be less obvious.

Previously, a group in Texas tested the variability of different instrumentation for fluorescence spectroscopy. They looked at the variability among spectra of three different standard trays that were recorded three times a day at three different locations using three spectrometers and four fiber optic probes, all having the same basic geometry. [5] Most of the variability was explained by differences in the spectrometer (70%) or fiber optic probe (15%). These investigators failed to assess differences due to the CCDs, which we believe can also affect Raman spectra.

Recently in our lab, a study has been done to assess the variability among spectra using different probe-based instrumentation systems. The probe, camera, and laser were varied. Using 8 different instrumentation combinations the spectral variance of skin spectra was estimated. The index finger of a volunteer was measured three times with each instrumentation combination. Each probe delivered 80 mW of 785 nm incident light onto the tissue and collected the reflected light. The integration time for each sample was held constant at three seconds. For each measurement, all overhead fluorescent lights were turned off and the sample was covered with an opaque cloth. Spectra from this set of experiments are shown in Figure A1.4, where the probe and spectrometer are constant but the camera and the laser vary.

In this figure, there is little difference in the two cameras used (Pixis and Andor). The variance was calculated and was found to be minimal when comparing the different cameras. The biggest variable was changing the probe (both had different geometries), but if the collection optics are kept the same, the spectra may be reliably compared.

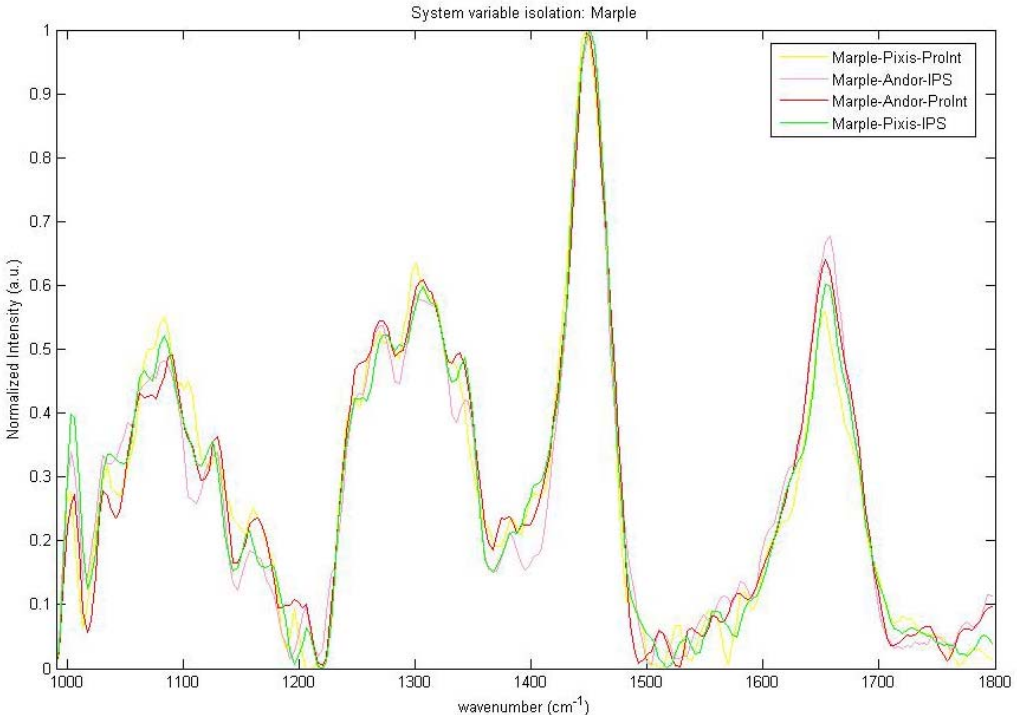


Figure A1.4: Average spectra after probe and spectrometer variable isolation.

A major limitation in this study is that only newer cameras were used. Both the Andor and the Pixis camera used in these experiments are TE cooled back-illuminated CCDs. The same chip made by the same company is in both of these cameras and they are only a few years old. Therefore, even though this study suggests that camera

variations are non-significant, using an older camera with a different chip might affect the spectra more significantly. Additionally, the spectra used in this analysis had significant post-processing, including white-light correction. We have used older, LN cameras in previous studies; however, the appropriate standards were not taken and the camera is currently broken. At this time, spectral differences resulting from chip type and age have not been evaluated.

In this chapter, differences due to instrumentation are discussed. Other sources of variations such as the doctor's experience and the probe slippage are also discussed.

A1.3 Patient Population Studied

Patient population data has been obtained over the past 8 years by either Elizabeth Kanter (EK) or Amy Robichaux-Viehoever (ARV), the PhD student who worked on the project before EK. Both collected measurements on women with evidence of cervical dysplasia; ARV collected mostly data from patients with HGSIL and hysterectomy patients. And EK collected data focusing on patients with LGSIL and Pap smear patients. Both ARV and EK have spectral data from healthy cervical tissue to serve of a control (either hysterectomy or Pap smear). Only the hysterectomy patients are included in this analysis because the Pap smear patient procedure was not consistent with that of the dysplasia study because no acetic acid was applied to the cervix before Raman measurements were taken. All of the measurements taken by ARV were taken using the LN cooled CCD, while all the measurements taken by EK were taken using the TE cooled CCD.

A1.3.1 Clinical Study Design - Dysplasia Patients

Eighty-five (33 AVR and 52 EK) patients undergoing either a colposcopy guided biopsy or a LEEP were recruited to participate in the study, as approved by the Institutional Review Board (IRB). To be eligible for enrollment, the patient must be undergoing a colposcopy guided biopsy or a LEEP, be between the ages of 18-75, and still have a cervix (no hysterectomy). Informed consent was obtained from each patient prior to the procedure. The cervix was then exposed and visually examined by the doctor. Acetic acid was applied to the cervix (to enable visualization of abnormal areas by turning them white), which was followed by iodine (to clean the tissue and show the location of squamous epithelium). The abnormal tissue was removed and placed in a fixative solution for pathology examination. Spectra were measured after the application of the acetic acid and before the application of the iodine. Spectra were acquired from multiple areas of tissue to be removed by the procedure, and 2 visually normal areas. Tissue that is removed was underwent pathological analysis.

A1.3.2 Clinical Study Design - Hysterectomy Patients

Thirty-six (33 ARV and 3 EK) patients undergoing hysterectomy were recruited to participate in the study, as approved by Vanderbilt's IRB. To be eligible for enrollment, the patient must have been undergoing a hysterectomy for something other than cervical disease and have been between the ages of 18-75. Informed consent was obtained from each patient prior to the procedure. The cervix was then exposed and visually examined by the doctor. Acetic acid was applied to the cervix to keep the procedure similar to that of the dysplasia patients. If the cervix was visually normal,

spectra were measured from multiple normal areas of tissue that were removed by the doctor during the procedure. Removed tissue underwent pathological analysis.

A1.4 Data Collection

Raman spectra were collected from multiple sites *in vivo* using a portable Raman spectroscopy system consisting of a 785 nm diode laser, a beam-steered fiber optic probe, an imaging spectrograph, and back-illuminated, deep-depletion, TE cooled CCD camera or a LN cooled CCD. All of these instruments are controlled with a laptop computer. Details of the system can be found elsewhere. [6] The fiberoptic probe delivered 80mW of incident light onto the tissue for 3-5 seconds. For each measurement, the overhead fluorescent lights were turned off.

Spectral calibration of the system was preformed each day using a neon-argon lamp, naphthalene, and acetaminophen standards to correct for system wavenumber, laser excitation, and throughput variations. The spectra were processed for fluorescence subtraction and noise smoothing using the modified- mean method; this method is described in detail elsewhere. [6] Following data processing, each spectrum was normalized to its mean spectral intensity across all Raman bands to account for variability in overall intensity. White light correction was not done because this data does not exist for the data collected by ARV. These normalized spectra were categorized according to histology as determined by the pathologist and by menopausal status before it was used for further comparison and analysis.

A1.5 Statistical Analysis

The same analysis technique was used as was explained in previous chapters. The process consists of two steps: the first, extraction of diagnostic features from the spectra using the nonlinear maximum representation and discrimination feature (MRDF), and the second, development of a probabilistic scheme of classification based on linear sparse multinomial logistic regression (SMLR) for classifying the nonlinear features into corresponding tissue categories. This technique was used because it was necessary to detect small differences in recorded spectra that were caused to the camera.

A1.6 Spectral results

The average of 3 Raman spectra from ARV normal cervix and EK normal cervix are shown in Figure A1.5. There are few visible differences; these could be due to processing and possibly camera resolution or quantum efficiency. The spectral resolution using the LN camera is believed to be around 10 wavenumber and the spectral resolution of the TE cooled camera is about 8 wavenumbers. The spectra taken with the LN (ARV) are much nosier than the spectra taken with the TE camera (EK). The spectra taken by ARV have a longer integration time of 5 seconds compared to the 3 seconds used for the EK data. This difference is due to the quantum efficiency of the cameras; the TE camera has higher quantum efficiency than the LN and therefore could produce smoother spectra in the same amount of time or even less. These differences could account for the more pronounced peaks in the EK normal spectra. In general, the shape is very similar and the same peaks are present.

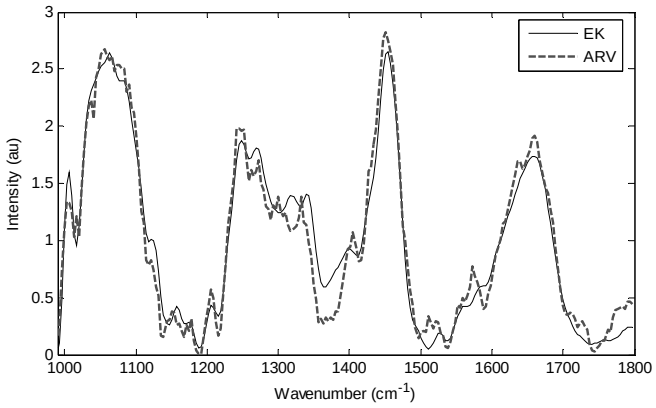


Figure A1.5. Average normal cervical spectra.

The average spectra from all of EK's data from HGSIL, LGSIL and normal are shown in Figure A1.6. When comparing differences between the camera and differences between pathology groups, they are on the same magnitude. Therefore, the differences in the camera need to be accounted for, and this is validated using statistical methods.

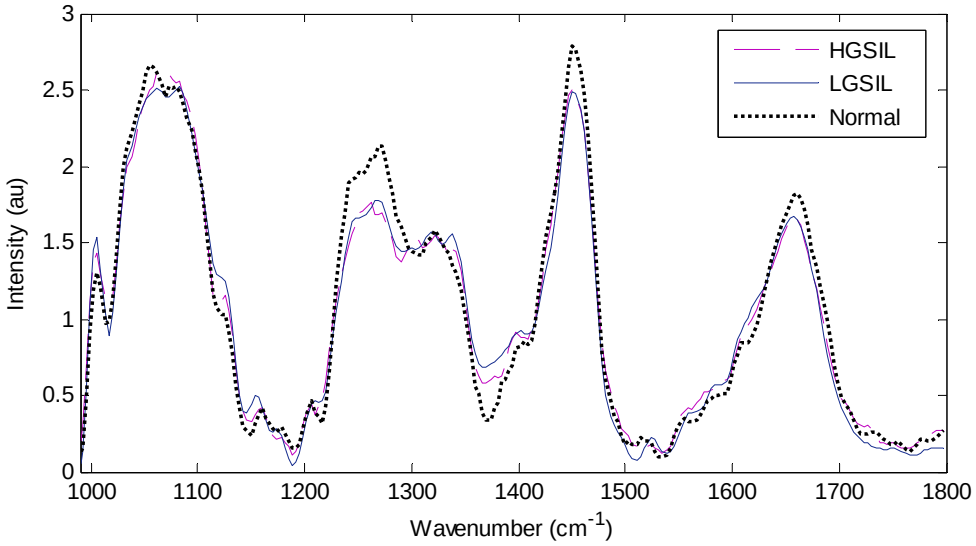


Figure A1.6. Average spectra from normal cervical, LGSIL and HGSIL.

A1.7 Classification results

A four- class algorithm (MRDF and SMLR) was used to classify the data (normal, metaplasia, LGSIL and HGSIL). In all cases, the classification results were obtained based on leave-one- sample- out cross validation of the entire data set. The classification was originally run 3 different times. The first time, the ARV and EK data were combined and the classification accuracy was 79.6%. The confusion matrix is shown in Table A1.1; it is shown in percentages for ease of understanding.

Table A1.1. Confusion matrix for all ARV and EK data. Classification accuracy 79.6%

		Histological Classification			
		Normal	Metaplasia	LGSIL	HGSIL
Raman Classification	Normal	85.8%	47.6%	25%	17.9%
	Metaplasia	1.7%	52.4%	0%	0%
	LGSIL	11.7%	0%	75%	0%
	HGSIL	0.8%	0%	0%	82.1%

Because the algorithm had such poor classification, the reason for misclassification was investigated. It was noted that most of the misclassified data was collected by ARV and not by EK. The dysplasia data was more heavily weighed on EK's data (it comprised of 61%) but the normal data was mostly ARV data; consequently, instrumentation variations are probably responsible for the misclassification and would explain why EK's data classified better.

Because the statistical analysis that is used throughout this dissertation is very complex, a much simpler method using PCA and LDA were used to see if instrumentation variations could be used to accurately separate the data. The all the normal data was run through PCA, and the first 14 PC's were used in order to use linear discriminate analysis. The results are shown in Figure A1.7.

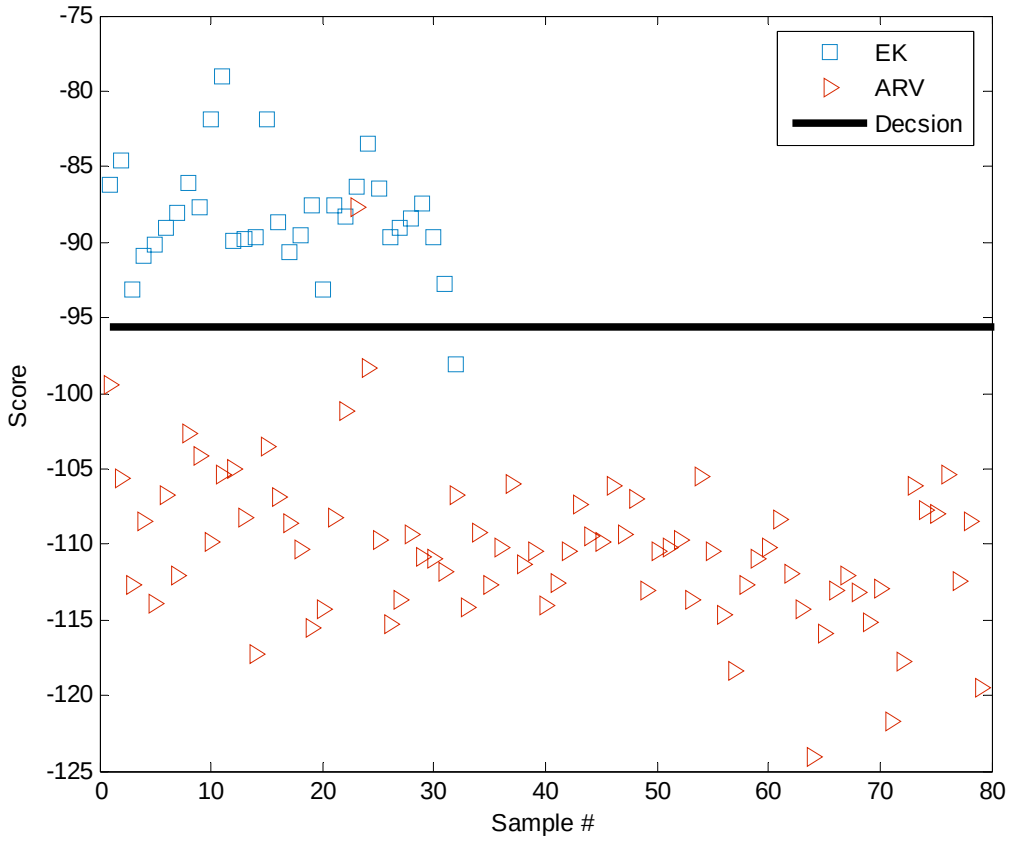


Figure A1.7. PCA/LDA results of the normal data classified by both systems.

This figure shows that there are significant differences between the two systems. The system used for EK was misclassified as the system used by ARV once, and the system used by ARV was classified as EK 1 time. This shows that the differences between the two systems are great enough that simple statistical methods can easily differentiate between the two systems, and this needs to be taken into consideration when using statistical algorithms.

In order to prove the system variations (specifically the changing of the camera – all other instrumentation remains unchanged) were responsible for poor classification, the algorithm was rerun for both ARV's data set and for EK's data set individually. The results are shown in Tables A1.2 and A1.3 respectively.

Table A1.2. Confusion matrix for all ARV data only. Classification accuracy 94%.

		Histological Classification			
		Normal	Metaplasia	LGSIL	HGSIL
Raman Classification	Normal	86.5%	0%	29%	16.6%
	Metaplasia	0%	100%	0%	0%
	LGSIL	10.8%	0%	71%	0%
	HGSIL	2.7%	0%	0%	83.4%

Table A1.3. Confusion matrix for all EK data only. Classification accuracy 87%.

		Histological Classification			
		Normal	Metaplasia	LGSIL	HGSIL
Raman Classification	Normal	86.5%	0%	13%	16.6%
	Metaplasia	0%	100%	0%	0%
	LGSIL	10.8%	0%	87%	0%
	HGSIL	2.7%	0%	0%	83.4%

ARV data classified correctly 94% of the time. The low grade was classified poorly in this particular algorithm because there was insufficient data. The biggest cause of concern 16.6% of HGSIL was classified as normal; this is a large number of false negatives and could be very dangerous because up to 20% of HGSIL may progress to invasive cancer if left untreated. [7]

When the algorithm was applied to EK's data, dysplastic cervix was classified with 87% accuracy. EK's data classification accuracy is less than ARV's classification accuracy, which could be attributable the EK's focus on LGSIL (LGSIL is much harder to classify). Its best performance was in metaplasia, where all the spectra were classified correctly, and was approximately 12-16% misclassified in the other 3 categories. In order to improve classification of normal, low grade, and high grade, the classification algorithm was modified to only include pre-menopausal women (as a consequence of the differences explained in Chapter 4). After all the data that was not from pre-menopausal women was eliminated and the classification algorithm was rerun. The classification

accuracy improved to 94% after making that correction and the confusion matrix is shown in Table A1.4.

Table A1.4. Confusion matrix for all EK pre-menopausal data only. Classification accuracy: 94%.

		Histological Classification			
		Normal	Metaplasia	LGSIL	HGSIL
Raman Classification	Normal	87.1%	0%	4.8%	0%
	Metaplasia	0%	100%	0%	0%
	LGSIL	12.9%	0%	95.2%	0%
	HGSIL	0%	0%	0%	100%

In this instance, the performance of metaplasia remained the same at 100%, but the classification of high grade improved from 83.4% to 100% and low grade improved from 87% to 95%. As shown in chapter 4, we can further improve this accuracy if we separate this data into before and after ovulation, but due to insufficient data, that has not yet been completed. A total of spectra from 5 patients were classified incorrectly, 2 spectra that were low grade were classified as normal, and 4 spectra that were normal classified as low grade. Of the normal spectra that were classified incorrectly, 2 of them were in patients in which the doctor was not 100% sure that the area that was being measured was completely normal; however it was not HGSIL like the majority of the cervix. Additionally, one of the others that was misclassified as normal, and one of the LGSIL that was misclassified as normal was taken on the doctor's first day of taking

measurements. The remaining patient (2 misclassified spectra) had no explanation of why the spectra might be misclassified unless it was a miscommunication between the operator and the doctor and the spectra actually classified correctly.

A1.8 Independent Validation

Independent validation is essential to testing any new clinical technique. Therefore, after the testing of the algorithm, a very small independent validation was done. We used spectra taken from patients undergoing a routine Pap smear that had come back abnormal. When the spectra were taken, the time of the measurement on the cervix was noted. The spectra were then run through the algorithm developed for this study. The colposcopy was done as normal, and the colposcopy results were compared to the algorithm results. The results of the six spectra that were run are as follows: 3 classified as LGSIL and 3 classified as normal. All of the spectra came from a cervix that the Pap smear suggested possible LGSIL. In all the areas that were classified as normal, there were no biopsies taken, but colpscopic examination suggested that these areas were indeed normal. In the area that classified as LGSIL, 2 of these regions were biopsied and were LGSIL, the other region was not biopsied but it was noted in the examination that it had a mosaic appearance. This region may have had some disease but there is no way to confirm this. Notably no acetic acid was used on these patients (which are why the Pap smear data was not included in the instrumentation variations), but, surprisingly, this does not seem to affect the outcome. All but one of the six spectra were classified correctly and the remaining spectrum does not have a matching diagnosis.

A1.9 Training

Probe slippage during measurements was an issue that may have affected the accuracy of readings taken during these studies. The doctor typically tells us that he or she thinks the probe slipped and that we should retake the measurements. It was found that when the probe slipped there was a characteristic shape associated of the spectra. Figure A1.8 shows the spectrum from the cervix of a single patient, one is a normal spectrum with no slipping, and the second is a spectrum that was recorded when the probe slipped. It has been shown that cervical spectra within the same patient are relatively similar, so visible differences like this are easy to observe. [2] A double peak in the 1000-1100 cm^{-1} wavenumber region is characteristic of this type of error, and the observance of a third peak is also common (the "third peak" in Figure A1.8 is more of a shoulder around 1120 cm^{-1}). The general shape of the slipped spectra is slightly different from that of the normal, but the relative intensity of the peaks is dramatically different. Another trend that was seen was much more subtle than the slipping phenomena, is that the doctors' first day of taking measurements is often misclassified. For example in table A1.4, most of the normal spectra that are misclassified are taken on the doctors' first day of taking measurements. This difference is much more subtle than the differences due to the probe slippage, because the graphical differences are not visible to the naked eye. It is possible that this could be due to the angling of the probe or not holding it flush to the tissue when taking a measurement. This issue suggests that before a doctor begins using this procedure in the clinic, a short training session needs to occur. It appears that after just a few measurements, the doctor can accurately hold the probe in the correct position and, therefore, the data classifies correctly.

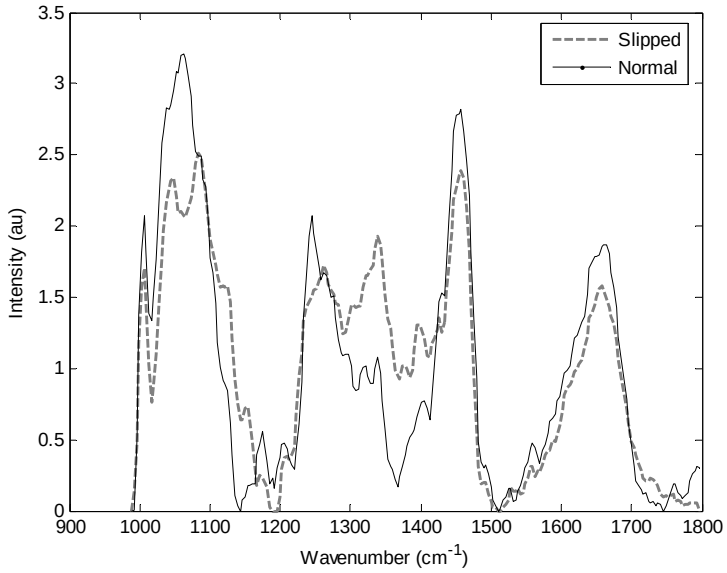


Figure A1.8: Raman spectrum from a normal cervix, the dotted line is from when the probe slipped and the solid line is from when it was held in place properly.

A1.10 Conclusions

The spectral differences that results from either camera variations or system variations are small and can not be easily seen. Nevertheless, these differences are significant enough to be picked up by the sophisticated algorithm employed by EK in this study. Also, it is necessary to either use the exact same equipment or do significant post-processing. Finally, physicians training with the proper use of the probes will increase the quality and accuracy of the data collected which will improve the accuracy of the patient's diagnosis.

A1.11 References

1. L. Knudsen, C. K. Johansson, P. A. Philipsen, M. Gniadecka and H. C. Wulf, "Natural variations and reproducibility of in vivo near-infrared Fourier transform Raman spectroscopy of normal human skin," *Journal of Raman Spectroscopy*, 33(7), 574-579, 2002.

2. A. Robichaux-Viehoever, "Raman Spectroscopy for in vivo, non-invasive Detection of Dysplasia of the Cervix," Ph.D. Biomedical Engineering, Vanderbilt University, 2003.

3. P. J. Caspers, G. W. Lucassen, R. Wolthuis, H. A. Bruining and G. J. Puppels, "In vitro and in vivo Raman spectroscopy of human skin," *Biospectroscopy*, 4(5), S31-S39, 1998.

4. B. W. Barry, H. G. M. Edwards and A. C. Williams, "Fourier-Transform Raman and Infrared Vibrational Study of Human Skin - Assignment of Spectral Bands," *Journal of Raman Spectroscopy*, 23(11), 641-645, 1992.

5. B. M. Pikkula, O. Shuhatovich, R. L. Price, D. M. Serachitopol, M. Follen, N. McKinnon, C. MacAulay, R. Richards-Kortum, J. S. Lee and E. N. Atkinson, "Instrumentation as a source of variability in the application of fluorescence spectroscopic devices for detecting cervical neoplasia," *Journal of Biomedical Optics*, 12(3), 2007.

6. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Applied Spectroscopy*, 57(11), 1363-1367, 2003.

7. American Medical Association, 1999.

APPENDIX 2

RAMAN MICROSPECTROSCOPY OF THE HUMAN CERVIX

A2.1 Introduction

A2.1.1 Normal Cervix

A typical cervix is 3-4 cm in length and approximately 2.5 cm in diameter, but the size and shape vary with age, parity, and hormonal status. [1] The cervix consists of two types of epithelia: the stratified non-keratinizing squamous epithelium that covers the ectocervix and is separated from the stroma by the basal layer, and the columnar epithelium that consists of a single layer of columnar cells and covers the surface of the endocervical canal. The interface of the two epithelia is called the squamocolumnar junction. The location of the squamocolumnar junction varies over a woman's lifetime due to age, hormonal status, child birth trauma, oral contraceptives, and pregnancy. The region where the columnar epithelium has been (or is being) replaced by squamous epithelium is called the transformation zone. [1] Almost all squamous cervical neoplasias begin at the functional junction and occur when the newly formed immature epithelium becomes atypical instead of normal mature squamous epithelium. [2]

A2.1.2 Cervical Disease and Progression

The term squamous intraepithelial lesion (SIL) refers to the development of neoplasia arising from the epithelium of the cervix and is often also referred to as cervical dysplasia. Clinically speaking, cervical lesions can be divided into low grade lesions

(LGSIL, or mild dysplasia) and high grade lesions (HGSIL, or moderate and severe dysplasia and carcinoma-in-situ (CIS)). This distinction is important, as patients with LGSIL lesions are typically followed but not treated, whereas patients with HGSIL are usually treated immediately with extended follow-up. [2-4] Most cervical lesions can be associated with a human papilloma virus (HPV) infection. HPV is a class of viruses that predominantly infect skin and mucosal membranes and produce characteristic epithelial proliferation, which can lead to malignant transformations. [2] There are several HPV strains that are more likely to induce this transformation; these include 16 and 18, which can be protected against using the Gardasil vaccine, and approximately 12 other subtypes.

A2.1.3 Relationship between Cervical Pathology and Cervical Biochemistry

As tissue becomes dysplastic, several changes occur, such as the nucleus to cytoplasm ratio increasing in the epithelium. This change is more drastic as the lesion becomes more serious and is one of the most important factors in assessing the grade of the disease. Another relevant change in the epithelium is a decrease in the amount of glycogen as disease progresses. [1] Changes associated with menopause also affect the biochemistry of the cervix, as discussed in Chapter 4. The epithelium thins, but the biochemistry of the tissue should remain constant but the ratio of components such as the epithelium and stroma should change. Other changes that occur in post menopausal women include the collagen matrix coalescing. [5]

A2.1.4 Raman Spectroscopy

As shown in the previous chapters, Raman spectroscopy has the potential to detect biochemical changes that occur in the cervix. DNA and other nuclear material, as well as glycogen, have characteristic Raman peaks, and changes in these peaks have been seen *in vivo*. In chapter 3, Figure 3.3 displays the average spectra taken *in vivo* from each of the following categories: LGSIL, HGSIL, normal, and metaplasia. In these spectra, the ratio of the 1260 cm^{-1} peak (associated with collagen) to the 1320 cm^{-1} peak (associated with DNA) increases in the LGSIL and HGSIL groups compared to benign cervix, meaning more DNA is present compared to the amount of collagen. Additionally, in chapter 4, Figure 4.2 displays the average spectra taken *in vivo* from each of the following categories: normal cervix from the 1^{st} part of the menstrual cycle, the 2^{nd} part of the menstrual cycle, peri-menopausal, and post menopause. In these spectra, the ratio of the 1260 cm^{-1} peak (associated with collagen) to the 1320 cm^{-1} peak (associated with DNA) increases in post-menopausal women compared to all other categories, meaning more signal is derived from the stroma compared to the epithelium since the epithelium is thinning (this change is small).

6.1.5 Optical Microspectroscopy

All of the differences described above were seen with probe-based measurements on *in vivo* tissue where the sampling volume is on the order of millimeters (approximately 1-2 mm). It is therefore important to understand the biochemical and morphologic basis of the spectral signatures, as well as the basis for differences in the signatures of the tissues when developing a diagnostic technique. Raman micro-

spectroscopy records spectroscopic information from tissue with high spectral and spatial resolution, and is thus a useful tool for understanding the biological basis of spectral-based tissue discrimination.

A2.1.6 Goal of this Study

The goal of this study was to determine from where a few of the differences that are observed in the *in vivo* Raman spectra originate, most the differences are expected to come from the epithelium in diseased patients since that is where the disease originates. Micro-Raman spectra were acquired from several different cervical tissue samples using a Renishaw Raman microscope. Stroma and epithelium of pre- and post-menopausal women and diseased and non-diseased women were characterized with this system. The spectra were compared to one another and to the probe data described in the pervious chapters. The majority of the signal that is seen in the probe data was collected from the stroma, and this proportion increases in post-menopausal women, likely due to the thinning of the epithelium.

A2.2 Methods

A2.2.1 Sample Collection and Preparation

Informed consent was obtained from patients undergoing LEEP excision and vaginal hysterectomy. Ethical approval for this study was obtained from Gloucestershire Local Ethics Committee (Institutional Review Board). Tissue samples, which were typically 1cm in diameter, were immediately snap frozen in liquid nitrogen upon

excision. For each patient, their age, menopausal status, and date of last menstrual period (if applicable) were noted.

Frozen sections with a thickness of 15-20 microns were placed onto calcium fluoride slides for Raman spectral mapping and stored frozen until measured. Serial seven micron sections from either side of the mapping section were obtained and stained with hematoxylin and eosin (H&E) for histopathological evaluation. An expert pathologist annotated each sample with the different pathologies, including normal ectocervix epithelium, normal endocervix epithelium, transition zone, and LGSIL or HGSIL. Glands, stroma, and warts due to HPV were also identified for correlation with frozen sections. A total of 12 samples were obtained from a total of 6 patients - 3 from hysterectomy patients and 3 from LEEP patients. Two of the LEEP patients had evidence of HPV (LGSIL) and HGSIL in other parts of the tissue that was not available to us. Two of the hysterectomy patients were post-menopausal, and the final one was pre-menopausal.

A2.2.2 Raman Microspectroscopy

Prior to Raman microspectroscopy, sections were thawed at room temperature. Raman spectra were acquired using a customized Renishaw System 1000 spectrometer with Streamline™ technology, a novel rapid technique for Raman data acquisition as described previously. [6] The system is shown in Figure A2.1.

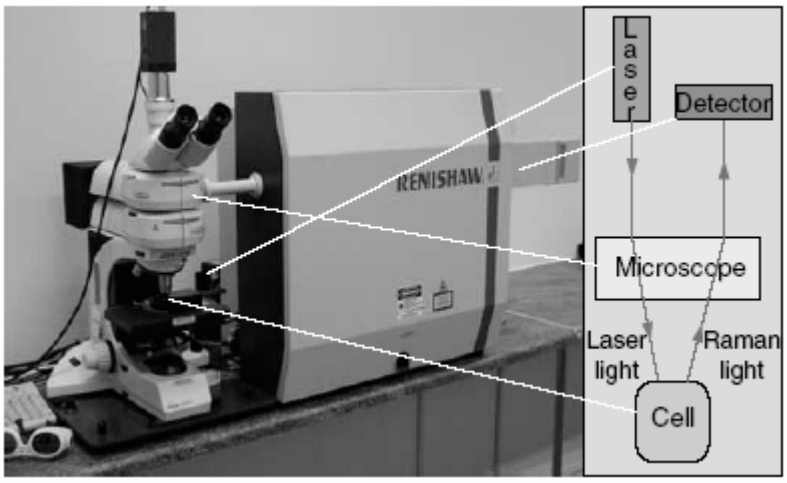


Figure A2.1: Schematic of the system used.

The system consists of two parts - a Renishaw spectrometer and a microscope. The spectrometer has a diode laser with a wavelength of 830 nm used for excitation and a 300 lines/mm grating used to disperse the scattered light, which was measured with a deep depletion charge coupled device (CCD) detector. The microscope consists of a Leica 50x long working distance objective used to focus the laser on the sample and to collect the Raman scattered light and a motorized translation stage. Renishaw Streamline proprietary hardware and software built in to WiRE™ 3.0 was used for Raman mapping.

Raman spectra acquired from regions of interest identified by the pathologist were acquired with a spectral acquisition time of 40 seconds and step size of 12.7 microns in both x and y directions. Typical maps were of the order of 50 x 100 spectra, which took approximately 14 hours to acquire. Some lager areas where acquired, and these maps took 3 days to acquire; fortunately, using the streamline this time was reduced from 3

weeks. A white light montage of each tissue sample was taken to correlate the Raman spectra to the histology image.

6.2.3 Data Processing

Standards where taken at the beginning of each measurement day, or the beginning of the measurement if it took several days. The spectrometer was calibrated daily with a silicon standard, which had its center wavenumber set at 520.4 cm^{-1}. Data processing was carried out using Matlab and the PLS toolbox (Eigenvector Technologies, Manson, Washington, USA). First cosmic rays were manually removed prior to principal component analysis (PCA). Pseudocolor principal component (PC) score maps were then generated. The pseudocolor score maps were compared to the H&E sections and were then used to select regions of different tissue pathological classification within both the epithelium and stroma. From the selected region, mean spectra were calculated for each region, and background subtraction was achieved by iterative subtraction of a 5^{th} order polynomial.

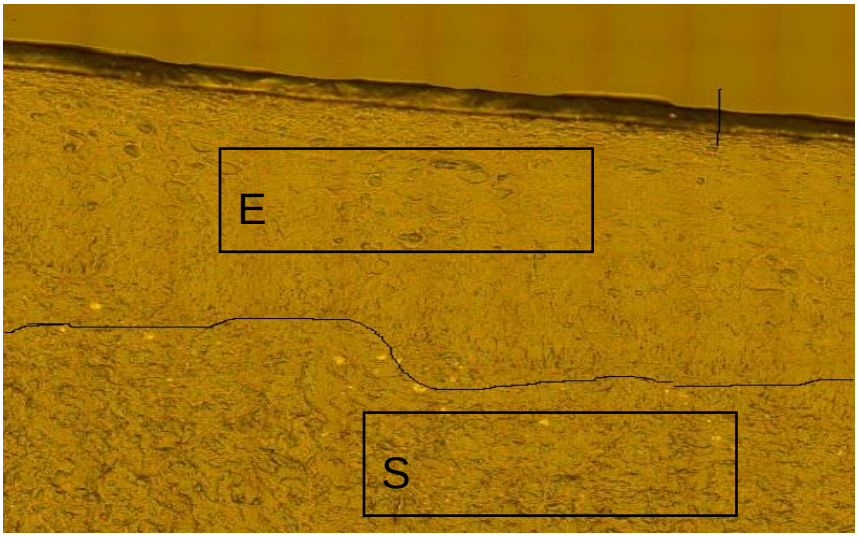
A2.2.4 Data Analysis

Raman peaks within the mean spectra that differed between the various pathology groups were identified and compared with the literature. Spectra were visually compared to spectra from the *in vivo* data. In some instances, the spectra were added together to best fit the spectra in the *in vivo* data to get the best visual approximation for where the majority of the signal originates. All resulting spectra were correlated to the existing data.

A2.3 Results

The white light image of a section of frozen normal cervical tissue is displayed in figure A2.2.a; regions of epithelium (E) and stroma (S) are marked. This image was mapped using Raman microspectroscopy, and the mean spectra from the epithelium and stroma are shown in figure 6.2.b. There are several differences between the mean spectra: the epithelium has a higher signal at 1086, 1126, and 1334 cm^{-1}, while the stroma has a higher signal at 1246, 1554, and 1614cm^{-1}.

(a)



(b)

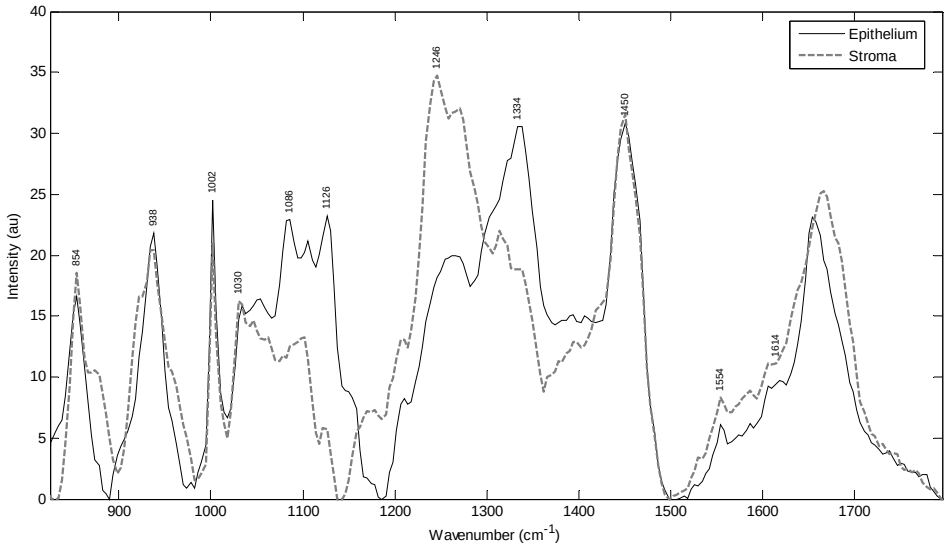


Figure 6.2: (a) White light image of the ecto-cervix; the bottom portion is epithelium, and the upper portion is stroma. (b) The corresponding average Raman spectra from each region.

In order to relate this result to the work discussed in previous chapters, these measurements taken with the microspectroscopy system were compared with those taken with a fiber optic probe-based system. The amount of signal that that comes from the relatively thin surface epithelium versus how much of the signal is due to the stroma was quantified. It was found that the best approximation of probe-measured spectra consisted of approximately 75% stroma and 25% epithelium using best fit and laying the two figures on top of one another. Figure 6.3 shows the combination of 75% stroma and 25% epithelium in the top panel and the spectra from a probe measurement in the bottom panel. Although this is an approximation (a more thorough approximation was not done because of the differences found in Appendix 1), the region from 1000cm^{-1} to 1150cm^{-1} is different between the two panels. This could be due to several factors; first, the optics are different in both of the set-ups and therefore different backgrounds might be present. As was shown in Appendix 1, spectra from two systems, especially if the geometries are different cannot be directly compared but crude approximations can still be made. Some of the differences are, silica fibers are used in the probe but not the micro-spectrometer system, and silica has a Raman peak at approximately 1000 cm^{-1}. Differences are also seen in the 1200-1400 cm^{-1} region, which could be due to a background subtraction issue. Additionally, a 785nm laser was used for the probe measurements, and an 830nm laser was used for the mapping measurements.

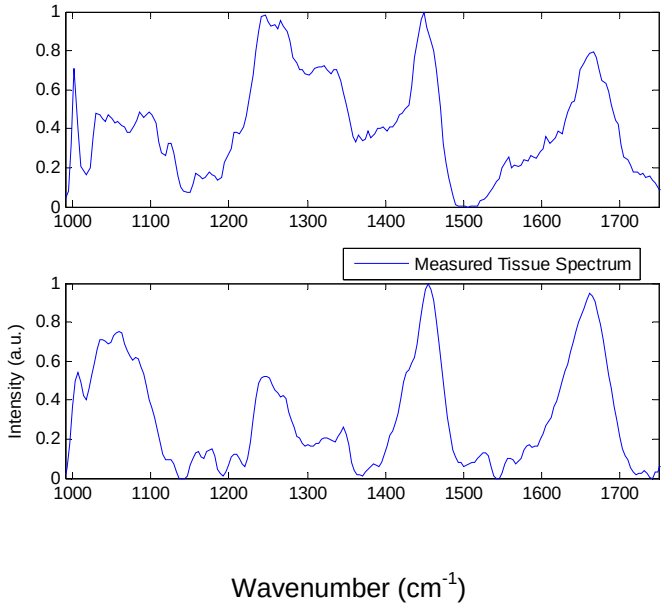


Figure A2.3. Raman spectra of 75% stroma and 25% epithelium in the top panel and the spectra from a probe measurement in the bottom panel

Average spectra of the stroma from pre (before ovulation)- and post-menopausal women are shown in figure A2.4. There are increases at 1030, 1098, 1314, 1126, 1314, 1558, and 1578 cm^{-1} in the pre-menopausal women compared to the postmenopausal women. Differences are expected because the cervix changes as a women ages. The average spectra from the epithelium in pre- and post-menopausal women is not shown because there were no visible differences between the spectra. The spectra in figure 4.2, which shows average spectra taken *in vivo* from normal cervix from the 1st part of the menstrual cycle, the 2nd part of the menstrual cycle, peri-menopausal, and post menopause, have some similar trends compared to the spectra in figure A2.4. The peak in the 1314 cm^{-1} to 1330 cm^{-1} range is higher in the pre-menopausal spectra compared to the post menopausal spectra, and the post-menopausal spectra are slightly higher at the

1450 cm^{-1} peak than the pre-menopausal. The epithelium might contribute to some of the changes that occur in the spectra shown in chapter 4, but since the spectra shown in 6.4 are from a woman who only underwent menopause in the last 5 years, the changes associated with menopause might not have taken full effect.

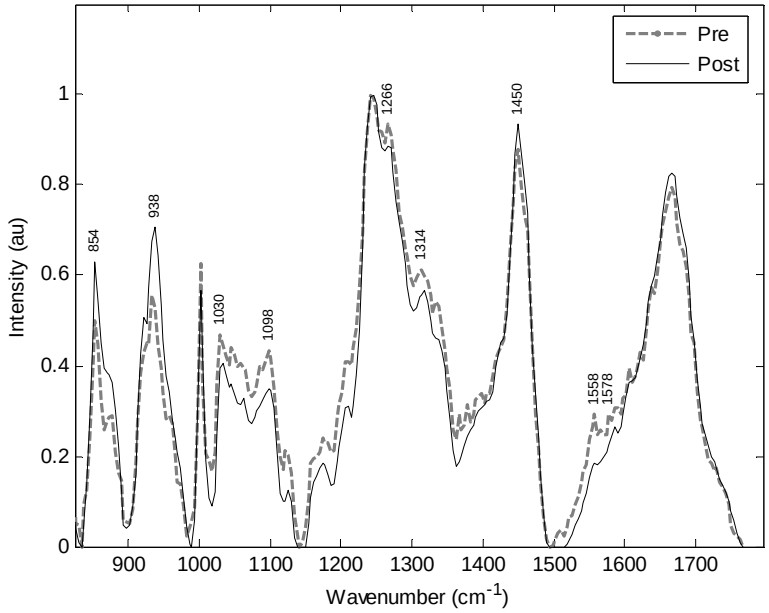


Figure A2.4: Average Raman spectra from the stroma of pre- and post-menopausal women.

Raman spectra were acquired from a sample that has both normal ectocervix and LGSIL (or mild HPV changes), a histology image of which is shown in figure A2.5. The region from which measurements were taken are highlighted and labeled (a-normal ectocervix, b-LGSIL, c-stroma under normal and d-stroma under LGSIL). Average spectra from normal ectocervix and LGSIL regions are shown in figure A2.6. Figure

A2.7 shows the average spectra from the stroma from a region under normal epithelium and the stroma from diseased epithelium.

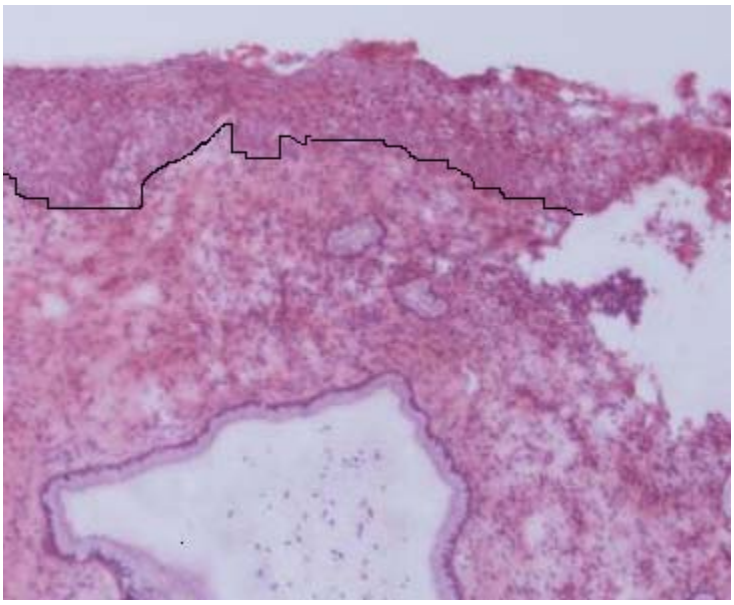


Figure A2.5: Histology image from a cervix that has evidence of LGSIL and normal. a-normal ectocervix, b-LGSIL, c-stroma under normal and d-stroma under LGSIL.

LGSIL, in this case HPV changes, is shown along with normal to highlight the different peak intensities as determined by an expert pathologist and based on the pathology of the rest of the tissue. The differences between normal epithelium in the cervix and HPV changes within the same sample are subtle. Slight increases in intensity can be seen at 1334 and 1082cm^{-1}, both of which are usually associated with DNA.

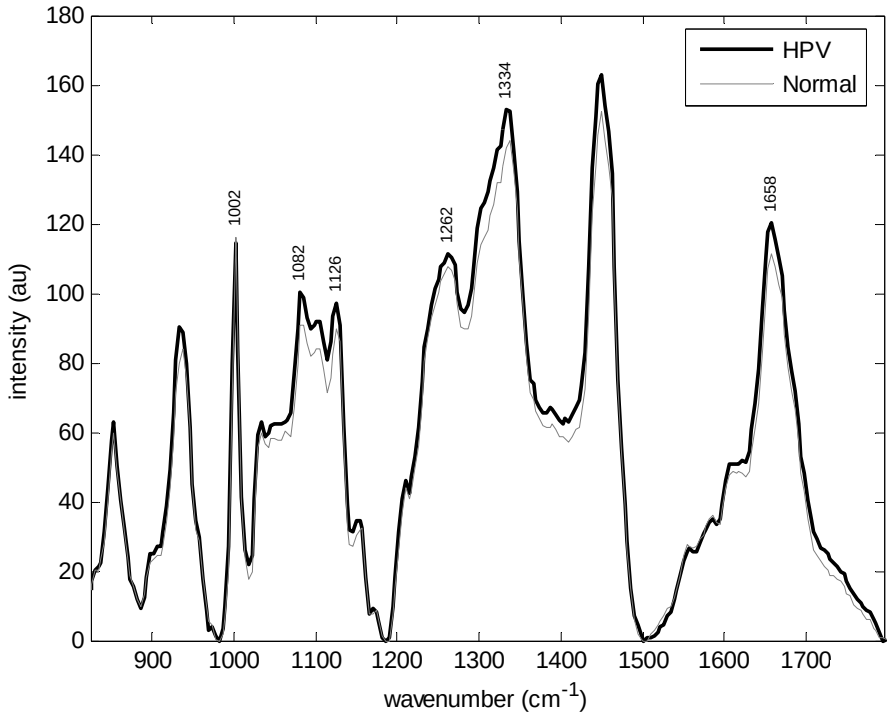


Figure A2.6: Average Raman spectra from the epithelium of HPV infected tissue and from normal cervical tissue.

There are larger differences between the stroma under HPV infected tissue and the stroma under normal tissue. The average spectra from the stroma of HPV infected tissue and from normal tissue, as depicted in figure A2.5, are shown in figure A2.7. There are increases in peak intensity from 1000 cm^{-1} to 1100 cm^{-1} and from 1250 cm^{-1} to 1350 cm^{-1}. The stroma under normal tissue seems to be more active than the stroma under the infected cervical tissue. This suggests that having disease in the epithelium does affect the underlying tissue. It is expected that in diseased tissue, the increase in signal associated with disease in the epithelium and thickening of the epithelium will increase the percentage of the signal coming form the diseased portion of the tissue, making differentiation of disease and normal possible.

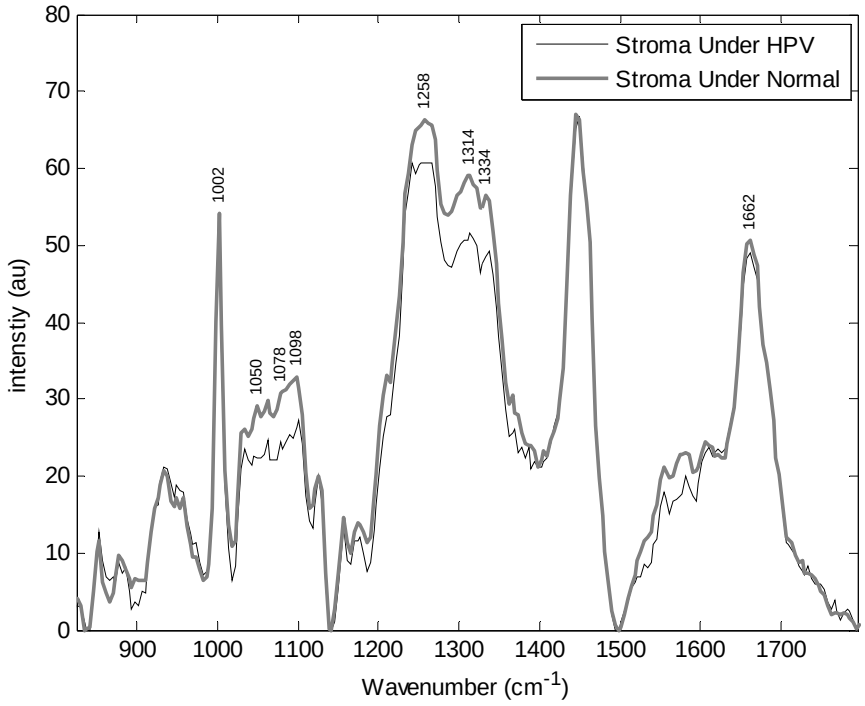


Figure A2.7: Average Raman spectra from the stroma of HPV infected tissue and from normal cervical tissue.

A2.4 Discussion

Human cervical tissue was investigated with a Raman micro-spectrometer microscope. This allowed for the differences between the epithelium and the stroma, the changes that occur with aging, and the differences that occur with disease to be highlighted. Many studies have been completed on *in vivo* cervical tissue, and have clarified where some of the changes in tissue originate. [7, 8], [Chapter 3-5]

The stroma is composed mostly of collagenous connective tissue, whereas the epithelium is a more active cellular layer. These differences can be seen when comparing the average Raman spectra from stroma and epithelium in Figure A2.1.b. The stroma has higher signal in regions that are usually associated with collagen, such as the 1002, 1266,

and 1555 cm^{-1} peaks. [9] Conversely, the epithelium has higher signal in regions associated with DNA, such as 1334 cm^{-1} [10], and lipid content, like the 1086 [11] and 1126 cm^{-1} peaks. [12] The components displayed in figure A2.2.b were then added together to estimate the relative contribution of the epithelium and stroma in the spectra taken from the normal cervix with a probe *in vivo*. The best estimate is that 25% of the signal comes from the epithelium and 75% of the signal comes from the stroma. This was somewhat unexpected since the probe geometry favors collection from the surface of a tissue, in this case the epithelium. The excitation photons can penetrate well into the stroma, however, since the epithelium is only about 200 microns thick.

The stroma is more active in pre-menopausal women than post-menopausal women because it is no longer needed in post-menopausal women. [1] Collagen is present in both the pre- and post-menopausal spectra, but there are some differences in the peaks associated with it. Overall, the trend is a decrease in the peaks associated with collagen (1030 cm^{-1} [13], 1126 cm^{-1}, 1314 cm^{-1} [14] and 1558 cm^{-1} [9]), but in two of the peaks associated with collagen (854 cm^{-1} [14] and 938 cm^{-1} [7]), there is an increase in relative intensity of collagen in menopausal women compared to pre-menopausal women. There is a structural change in the collagen in the stroma that occurs with menopause. The collagen matrix coalesces, which suggests a change in the dynamics of the matrix that results in the shrinking of the cervix [5], which explains why some of the peaks associated with collagen increase and others decrease. There is also an increase in a few of the peaks usually associated with DNA, such as 1098 cm^{-1}, 1314 cm^{-1}, and 1578 cm^{-1}. The decrease in DNA is expected because the cervix is primarily regulated by ovarian estrogens. The estrogen is present in significantly lower levels in the post-menopausal

women, and therefore it is expected that the cellular content, and thus DNA content, of the cervix decreases.

Several differences are expected between dysplasia and normal spectra (figure A2.6 and A2.7). Dysplasia only directly affects the surface epithelium. The grading of cervical dysplasia is dependant on the degree of epithelium involvement; as soon as disease invades into the stroma, it is considered cervical cancer. In figure A2.6, differences between the epithelium of a normal vs. diseased cervix are displayed. This cervix has very mild disease (HPV changes), so it is expected that there will be some changes in the Raman originating from the epithelium. Some of the changes that occur in the spectra can be associated with DNA, in particular at 1334 cm^{-1}. [15] It is expected that as a tissue starts to become more active, and in this case more dysplastic, the amount of cellular proliferation and therefore the amount of DNA should increase. Although this change is very subtle in the example shown, this is expected because there is a minimal amount of disease present. In fact, this type of disease will regress back to normal the majority of the time and is very unlikely to develop into anything serious.

The changes that are seen in the average Raman spectra from the stroma below HPV infected tissue and from the stroma below normal tissue, in figure A2.7, are more unexpected than what is seen in figure A2.6. In this graph (figure A2.7), there are increases in two regions of the spectra of stroma below a normal compared to stroma below a disease. This suggests that disease could actually indirectly affect the underlying cervix. There are some differences in peaks, such as 1316 cm^{-1} [16] and 1334 cm^{-1} [15], that are associated with DNA, and increases in peaks that are associated with glycogen – 1048 cm^{-1}, 1083 cm^{-1}, 1256 cm^{-1} and 1333 cm^{-1} [13]. It is known that

the glycogen in the epithelium decreases with disease [1], but there is no mention of what happens in the stroma as a result of disease in the epithelium. Other increases are in the lipid content of the cervix. The peaks at 1260 cm^{-1} to 1304 cm^{-1} [16] are associated with the Amide III band. All of these increases may suggest healthy tissue, but the reason for this is unknown. It could be due to the nutrients in the blood being used to produce new dysplastic cells, or it could be that disease decreases the activeness of the surrounding tissue. This phenomenon needs to be studied in more detail.

The findings in this chapter will help improve the detection abilities of Raman spectroscopy *in vivo*. It cements the need to separate pre- and post menopausal women, since the thinning of the epithelium increases the stromal contribution, and there are differences in the normal stroma between these two groups. It also suggests that small changes in the cervix associated with the disease can be seen in both the epithelium and the stroma, so the collection volume is reasonable for cervical dysplasia detection.

A2.5 Conclusions

We were able to show that using Raman mapping of the human cervix, one can see differences between the epithelium and stroma, and that when this is compared to probe data, the stroma dominates the signal even though it is subsurface. It was also found that Raman can detect changes associated with menopause in the stroma but not in the epithelium. In a previous chapter (4), it was demonstrated that using the probe-based system, there were statistical differences between pre- and post-menopausal women. It was found that these differences are due to changes in the stroma due to the coalescence

of collagen. Additionally, it was found that HPV affects the spectra of both the epithelium and the stroma.

A2.6 Future directions and limitations

Several experiments need to be done to make this study more complete. First, all of the mapping data was done at 830nm, whereas all the probe data was done at 785nm, so a comparison of the two different wavelengths needs to be done. Ideally we need to do the comparison on the same piece of tissue. Also, high grade dysplasia and cancer need to be added to the dysplasia part of the study. Unfortunately, obtaining this tissue is very difficult. It is almost impossible to obtain dysplasia tissue in the US; therefore, we have been working with a group in England to obtain these tissue samples. We have been collecting tissue for several months and have not been able to obtain any high grade samples and only 2 low grade samples.

A2.7 References

1. J. W. Sellors and R. Sankaranarayanan, *Colpsocopy and Treatment of Cervical Intreaepithelial Neoplasia: A Beginners' Manual*, International Agency for Research on Cancer, Lyon, 2003.

2. T. Wright, R. J. Kurman and A. Ferenczy, "Cervical Intraepithelial Neoplasia," *Blaustien's Pathology of the Female Genital Tract*, Springer-Verlag, New York, 1994.

3. J. L. H. Evers and M. J. Heineman, *Gynecology - A Clinical Atlas*, CV Mosby Co, St.Louis, 1990.

4. I. Ramzy, *Essentials of Gynecologic and Obstetric Pathology*., Appleton - Century - Crofts, Norwalk, 1983.

5. J. Lorrain, *Comprehensive Management of Menopause*, Springer-Verlag, 1994.

6. J. Hutchings, C. Kendall, N. Shepherd, H. Barr, B. Smith and N. Stone, "Rapid Raman microscopic imaging for potential histological screening", *Biomedical Optical Spectroscopy*, SPIE, 6853(685305-9), 2008.

7. A. Mahadevan-Jansen and R. Richards-Kortum, "Raman Spectroscopy for the detection of cancers and precancers," *J Biomed Opt*, 1(1), 31-70, 1996.

8. A. Robichaux-Viehoever, E. Kanter, H. Shappell, D. Billheimer, H. Jones and A. Mahadevan-Jansen, "Characterization of Raman spectra measured in vivo for the detection of cervical dysplasia," *Appl Spectrosc*, 61(9), 986-993, 2007.

9. C. J. Frank, D. C. B. Redd, T. S. Gansler and R. L. McCreery, "Characterization of Human Breast Biopsy Specimens with near-Ir Raman-Spectroscopy," *Analytical Chemistry*, 66(3), 319-326, 1994.

10. E. B. Hanlon, R. Manoharan, T. W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari and M. S. Feld, "Prospects for in vivo Raman spectroscopy," *Physics in Medicine and Biology*, 45(2), R1-R59, 2000.

11. Y. Ozaki and A. Mizuno, "Molecular Aging of Lens Crystallins and the Life Expectancy of the Animal - Age-Related Protein Structural-Changes Studied Insitu by Raman-Spectroscopy," *Biochimica Et Biophysica Acta*, 1121(3), 245-251, 1992.

12. P. J. Caspers, G. W. Lucassen, R. Wolthuis, H. A. Bruining and G. J. Puppels, "In vitro and in vivo Raman spectroscopy of human skin," *Biospectroscopy*, 4(5), S31-S39, 1998.

13. C. Kendall, "A study of Raman spectroscopy for the early detection and classification of malignancy in oesophageal tissue.," *Cranfield University*, PhD Thesis, 2002.

14. C. J. Frank, R. L. McCreery and D. C. B. Redd, "Raman-Spectroscopy of Normal and Diseased Human Breast Tissues," *Analytical Chemistry*, 67(5), 777-783, 1995.

15. A. Nijssen, T. C. B. Schut, F. Heule, P. J. Caspers, D. P. Hayes, M. H. A. Neumann and G. J. Puppels, "Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy," *Journal of Investigative Dermatology*, 119(1), 64-69, 2002.

16. A. Mahadevan-Jansen, W. F. Mitchell, N. Ramanujam, U. Utzinger and R. Richards-Kortum, "Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue in vivo," *Photochemistry and Photobiology*, 68(3), 427-431, 1998.