NEAR-INFRARED RAMAN SPECTROSCOPY FOR EARLY DETECTION OF CERVICAL PRECANCER

MO JIANHUA

NUS

NATIONAL UNIVERSITY OF SINGAPORE
2010

NEAR-INFRARED RAMAN SPECTROSCOPY FOR EARLY DETECTION OF CERVICAL PRECANCER

MO JIANHUA

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY OF ENGINEERING

DIVISION OF BIOENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2010

Acknowledgements

I would like to express my deep appreciation to my advisor, Dr. Zhiwei Huang, for his professional guidance, unending encouragement and great patience as well as financial support during the course of my PhD candidature in the past 5 years. He taught me how to do scientific research in the area of biomedical optics from the research in laboratory to clinical trials, including experiment design, data processing and analysis, and scientific research article writing. I am very sure what Dr. Huang taught me will not only help me to complete my PhD study but also benefit my future career forever.

I would be grateful to Dr. Arunachalam Ilancheran, Dr. Jeffrey Low Jen Hui and Dr. Ng Soon Yau Joseph from Department of Obstetrics and Gynaecology, National University Hospital, Singapore. They offered me a great help in conducting the clinical trials by showing great patience and also taught me much medicine knowledge related to my thesis work. I would also appreciate Dr. Si-shen Feng and Dr. Nanguang Chen for their kind advice and time on my research work.

I would express my gratitude to Dr. Wei Zheng, Dr. Franck Jaillon, Sengknoon Teh for their kind help on my thesis work. I would thank other members of Dr. Huang's group, including Mads Bergholt, Shiyamala Duraipandian, Fake Lu, Hao Li, Jian Lin, Kan Lin, Xiaozhuo Shao, Sathish Kumar Sivagurunathan, Clement Yuen, Hamed Zaribafzadeh. I would also thank other students and staffs in optical bioimaging laboratory in National University of Singapore in Singapore. They are Ling Chen, Shaupoh Chong, Shanshan Kou, Linbo Liu, Weirong Mo, Weiteng Tang, CheeHowe Wong and Qiang Zhang. I spent happy time with all of them above during my thesis study in optical bioimaging laboratory.

To my girl friend, Yanan Li, thank you for staying with me and for your warm encouragement and concern in the hardest time of my thesis work.

To my families, my parents, brother and sister-in-law, you always show great material and moral support to me which serves as the pivot of my research and study. Hope that the completion of my thesis can be a good return for you. I will try my best to make you proud forever.

I would like to acknowledge the financial support to my research from the following

funding agencies: the Academic Research Fund from the Ministry of Education, the Biomedical Research Council, the National Medical Research Council, and the Faculty Research Fund from National University of Singapore.

Table of Contents

Abstract

This thesis work was centered on detecting cervical precancer using near-infrared (NIR) Raman spectroscopy. A rapid and portable NIR Raman spectroscopy system coupled with a specially designed ball lens fiber-optic Raman probe was successfully developed for *in vivo* tissue diagnosis and characterization. Firstly, Raman measurement was conducted on biopsied cervical tissues to test the feasibility of NIR Raman spectroscopy for the detection of cervical precancer. A good classification with an accuracy of 92.5% between benign and dysplasia (i.e., LGSILs and HGSILs) tissues was achieved *ex vivo*, encouraging the extension of our *ex vivo* work to *in vivo* study. Monte Carlo simulation method was employed to evaluate the performance (i.e., collection efficiency and depth-selectivity) of the ball lens fiber-optic Raman probe designs with various configurations (i.e., the diameter and refractive index of the ball lens). We demonstrated that the ball-lens NIR Raman spectroscopy developed is able to acquire good-quality Raman spectra of cervix *in vivo*. We demonstrated for the first time that NIR Raman spectroscopy in the high wavenumber (HW) region has the potential for the diagnosis of cervical precancer using our in-house developed Raman system and exhibits comparable diagnostic performance as Raman spectroscopy in fingerprint region. We also demonstrated that combining NIR autofluorescence and Raman spectroscopy can further improve the diagnosis of cervical precancer. We also evaluated the performance of ultraviolet/visible autofluorescence and diffuse reflectance spectroscopy in the characterization of cervical dysplasia and finally combined them with NIR Raman spectroscopy. It was found that optimal diagnosis of cervical precancer could be achieved by combining all these three different spectroscopic techniques together. The work completed in this thesis promotes some future directions to further optimize the diagnosis and detection of cervical precancer

in vivo using Raman spectroscopy. One of the major directions is to develop robust software integrated with Raman spectral data preprocessing, statistical modeling for real-time *in vivo* tissue diagnosis and characterization. Another major direction is to develop fluorescence image-guided Raman spectroscopic diagnosis system to further facilitate and improve early diagnosis and detection of cervical precancer in clinical settings.

List of Figures

List of Tables

List of Abbreviations

AF	=	autofluorescence
ANOVA	=	analysis of variance
ANSI	=	American National Standards Institute
BCC	=	basal cell carcinoma
BFL	=	back focal length
BP	=	band-pass
BI-CCD	=	back-illuminated CCD
CART	=	classification and regression tree
CCAM	=	congenital cystic adenomatoid malformation
CCD	=	charge-coupled device
CIN	=	cervical intraepithelial neoplasia
CIS	=	carcinoma in situ
CPC	=	compound parabolic concentrator
DNA	=	deoxyribonucleic acid
EEMs	=	excitation-emission matrics
EFL	=	effective focal length
FAD	=	flavin adenine dinucleotide
FDA	=	fisher discriminant analysis
FI-CCD	=	front-illuminated CCD
FT	=	frourier transform
FTIR	=	frourier transform infrared spectroscopy
FWHM	=	full width at half maximum
GI	=	gastro intestine
HCA	=	hierarchical cluster analysis
HGSILs	=	high grade squamous intraepithelial lesions
HP	=	*Helicobacter-pylori*
HPV	=	human papillomavirus
HW	=	high wavenumber
IM	=	intestinal metaplasia
IR	=	infrared
KCA	=	K-means cluster analysis
LD	=	linear discriminantion
LDA	=	linear discriminant analysis
LGSILs	=	low grade squamous intraepithelial lesions
LP	=	long-pass
MC	=	Monte Carlo
MDC	=	multispectral digital colposcopy
NA	=	numerical aperture
NADH	=	reduced form of nicotinamide adenine dinucleotide
NIR	=	near-infrared

os	=	orifice of the uterus
pap	=	papanicolaou
PC	=	principal component
PCA	=	principal component analysis
RNA	=	ribonucleic acid
ROC	=	receiver operating characteristic
SCC	=	squamous cell carcinoma
SD	=	standard deviation
SILs	=	squamous intraepithelial lesions
SMA	=	SubMiniature version A
SNR	=	signal to noise ratio
SVM	=	support vector machine
TCC	=	translational cell carcinoma
TE	=	thermoelectric
UV	=	ultraviolet
VIA	=	visual inspection after applying acetic acid
VIS	=	visible
VILI	=	visual inspection after applying Lugol's iodine
VUS	=	volume under the three-way ROC surface

Chapter 1 Introduction

1.1 Overview

Cancer is a major cause of human death. As the introduction of various prevention programs, the incidence and mortality rates declined in recent decades. However, the number of cancer and cancer-induced deaths is still increasing, which may be explained by the increased size of worldwide population and the population ageing. In particular for cervical cancer, it is the second most frequent cancer among women and it is estimated that there are 493,243 new cervical cancer cases in 2010 [1]. Meanwhile, there are 273,505 women dying from cervical cancer, which are the third most frequent cancer-induced deaths following breast (410,712) and lung (330,786) among women [1]. Therefore, great efforts are desired to prevent cervical cancer.

The prevention of cervical cancer usually comprises three procedures, including screening, diagnosis and treatment. The correct screening and diagnosis play a key role in the prevention of cervical cancer. At present, papanicolaou (pap) smear screening coupled with colposcopic diagnosis is the most common method for the prevention of cervical cancer. Pap smear can yield a sensitivity and specificity of around 60% in the detection of cervical precancer [2]. Colposcopic examination can improve the sensitivity to be above 90%; however, its specificity is even worse (~40%) [3]. Moreover, histopathology remains the gold standard for precancer and cancer diagnosis, which requires invasive biopsy, lengthens the diagnostic period and increases the cost. In this situation, optical spectroscopic technique has recently emerged as a promising technique to aid in the prevention of cancer by showing advantages of noninvasive, real-time and high-accuracy screening/diagnosis. Till now, the common spectroscopy used for screening/diagnosing precancer and cancer in the

cervix includes fluorescence, reflectance, infrared and Raman spectroscopy. They characterize various tissue pathologies by probing the changes in morphology, biochemical composition of tissue associated with tissue maligancy. In particular for Raman spectroscopy, it is a vibrational spectroscopic technique and molecular fingerprint probe, and has been applied for the detection of cancer and precancer in various human organs.

In this study, we aimed to explore the potential of NIR Raman spectroscopy in both fingerprint and high wavenumber (HW) regions for the *ex vivo* and *in vivo* detection of cervical precancer. To further enhance the acquisition of Raman signal originating from the epithelium of cervical tissues, a fiber-optic Raman probe coupled with a ball lens was designed and evaluated by using Monte Carlo (MC) simulation method. We also investigated the feasibility of combining NIR autofluorescence (AF) and Raman to improve the diagnosis of cervical precancer. In addition, we evaluated the performance of different optical spectroscopic techniques (i.e., NIR Raman, ultraviolet/visible (UV/VIS) autofluorescence and reflectance spectroscopy) in the detection of cervical precancer *ex vivo*; meanwhile, we studied if the diagnosis can be improved through combining the three methods as compared to either of them alone.

In this chapter, firstly, we will introduce the Raman effect and Raman spectroscopy instrumentation. Then, we will review the work on Raman spectroscopic diagnosis of various human cancers. Subsequently, we will present the background knowledge about cervical precancer and cancer. Next, we will review the work on the use of Raman spectroscopy and other alternative optical spectroscopic techniques (i.e., fluorescence, reflectance and infrared spectroscopy) for the detection of cervical precancer and cancer. Finally, we will present the motivations, objectives and organization of this thesis.

1.2 Raman Spectroscopy

1.2.1 The Raman Effect

The Raman effect was discovered by Chandrasekhara Venkata Raman in 1928, who was awarded the Nobel Prize in physics in 1930 for his work on discovering Raman scattering. The Raman scattering phenomenon was observed as an additional radiation to Rayleigh scattering (elastic scattering) and fluorescence emission when C. V. Raman was verifying the Rayleigh scattering. When light interacts with a molecule, a majority of photons are elastically scattered without frequency changes relative to the incident photons. This process is called Rayleigh scattering which is a classical theory of light scattering formulated by Lord Rayleigh in 1871. Meanwhile, a small fraction of light photons (approximately 1 in 10^8 incident photons) undergoes energy exchange with the molecule and consequently shows a frequency-shift against the incident photons. The process is defined as inelastic scattering and is also termed Raman scattering. In theory, the light interaction with a molecule leads to a polarization of the molecule and then the polarized molecule exhibits an induced dipole moment caused by the external field. The induced dipole moment P is proportional to the electric field E and to a property of the molecule called the polarizability α as shown in the following equation [4]:

$$P = \alpha E; \qquad E = E_0 \cos 2\pi \nu_0 t\,; \qquad P = \alpha E_0 \cos 2\pi \nu_0 t \tag{1.1}$$

where E_0 and ν_0 are the vibrational amplitude and frequency of the incident light, respectively. The polarizability α is dependent upon the position of the nuclei in the molecule. For a molecule containing N atoms, there are 3N degrees of freedom available to the nuclei. Of there, 3N-6 (3N-5 for a linear molecule) results in the vibrations of the molecule. Considering a diatomic molecule with the single normal coordinate Q_1, the induced dipole moment is as below [4]:

$$P = \alpha_0 E_0 \cos 2\pi \nu_0 t + \frac{1}{2} E_0 Q_1^0 \left(\frac{\partial \alpha}{\partial Q_1} \right)_0 \times \left[\cos 2\pi t \left(\nu_0 + \nu_1 \right) + \cos 2\pi t \left(\nu_0 - \nu_1 \right) \right] \quad (1.2)$$

where α_0 is the inherent polarizability of the molecule, Q_1^0 and ν_1 are the vibrational amplitude and frequency of the molecule, respectively. The first term represents Rayleigh scattering, and if $\partial\alpha/\partial Q_1$ is nonzero, Raman scattering occurs. The second and third terms represent anti-stokes and stokes Raman scattering, respectively, as shown in Fig. 1.1.

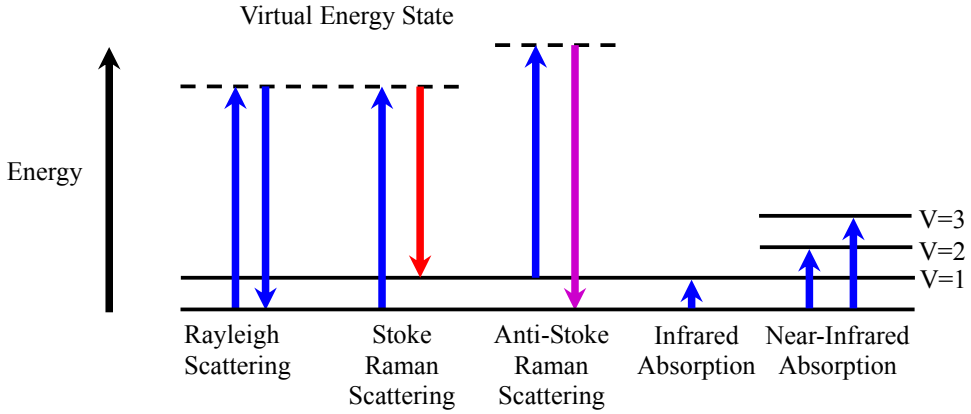


Figure 1.1 Energy transition diagram of vibrational spectroscopy. V is the vibrational quantum number.

For stokes Raman scattering, the scattered photon has a lower energy (longer wavelength) than the exciting photon. On the contrary, for anti-stokes Raman scattering, the scattered photon is located at shorter wavelength compared to the exciting photon. Conventional Raman spectroscopy is based on stokes Raman scattering. A Raman spectrum is created by determining the Raman intensity as a function of frequency shift ($1/\lambda_{excitation}$-$1/\lambda_{Raman}$), so called Raman shift which is quantified in wavenumber (cm^{-1}). Raman spectrum is characterized by a few distinct bands attributed to specific group of vibrational bonds in the molecules of the sample. Raman spectroscopy has proved to have the potential for diagnosing cancer and precancer through measuring Raman spectral changes representing the structural and

conformational changes of biomolecules associated with the cancerous transformation.

In addition, it is noticed in Fig. 1.1 that there exists another vibration spectroscopy (i.e., infrared (IR) or near-infrared (NIR) absorption spectroscopy). The extent of energy exchange during Raman scattering is identical to the energy absorbed during IR absorption. This implies that the frequency shift for certain vibration band of the same molecule remains the same for Raman scattering and IR absorption. However, the selection rule of Raman scattering differs from that of IR absorption. A molecule absorbs IR light only when the dipole moment changes during the molecular vibration. Whereas, the Raman effect is caused by an oscillation-induced dipole moment, which means that the molecular interaction with light is through the polarizability of the molecule. Therefore, not all the molecules are both Raman-active and IR-active, which makes Raman spectroscopy and IR spectroscopy complementary to each other.

1.2.2 Raman Instrumentation

Raman spectroscopy mainly consists of four parts, including excitation light source, spectrograph, detector, and sampling module. In principle, the light is delivered to the sample by the sampling module and then interrogates the sample. The Raman scattered photons in the sample undergo multi-scattering and absorption, and subsequently are collected by the sampling module. Then, the collected Raman photons are fed into the spectrograph and collected by the detector. Finally, spectrum is created with the output of the detector.

It is noted that the Raman shift of specific molecular vibration is independent of excitation light and Raman scattering is very weak. Therefore, high power monochromatic excitation light is required for Raman spectroscopy. As the invention and advance of laser technology, laser light from near-ultraviolet to near-infrared

regions (e.g., 488-, 515-, 785-, 830-, and 1064-nm) dominates the light source for Raman spectroscopy [4]. As for Raman spectrograph, it can be categorized into two types, including dispersive and non-dispersive. Dispersive spectrograph separates Raman photons spatially and then disperses them onto a multichannel detector using grating. In contrast, non-dispersive Raman spectrograph does not require spatial separation of Raman photons. At present, Fourier Transform (FT) Raman spectroscopy is one of the common non-dispersive Raman spectrograph forms. It is a multiplexing technique based on Michelson interferometer which modulates all the different wavelengths to produce a complex "interferogram". The interferogram is detected by a single-channel detector and eventually converted to spectrum by Fourier Transform. In the early time of Raman spectroscopy, FT Raman spectroscopy is the most prevalent. As the invention and advance of charge-coupled device (CCD), dispersive Raman spectroscopy based on CCD has become the major form of Raman spectroscopy.

In addition to light source, spectrograph, and detector, sampling module is also a key part of Raman spectroscopy and exerts a big impact on the sensitivity and application scope. For example, prior to the introduction of fibers to Raman spectroscopy, it is hard to achieve a remote control of Raman spectroscopic measurement. Till now, Raman sampling module has the following forms: (1) 90° or 180° scattering mode based on normal lens; (2) back-scattering or forward-scattering microscopic mode; (3) fiber-optic sampling using various fiber-optic probes.

1.2.3 Cancer Diagnosis by Raman Spectroscopy

1.2.3.1 Raman-active Biomolecules

Raman spectroscopic diagnosis of precancer and cancer is based on the fact that a big amount of molecules in biological tissues are Raman-active and meanwhile show

significant changes accompanying tissue premalignant and malignant transformation. Tissue Raman spectrum is a mixture of Raman signals from various molecules in tissue and consequently can represent the changes of those molecules. The spectrum variation among different tissue pathologies due to the molecular changes enables Raman spectroscopy to differentiate tissues under different histopathological statuses. It has been recognized that the major Raman-active biomolecules which are sensitive to tissue premalignant and malignant changes are proteins, lipids, nucleic acids and carbohydrates [5].

Proteins take up about 20% of total body weight, provide structural support and are also involved in all metabolic processes. Proteins and synthetic polypeptides consist of amino acids joined together by peptide bond (-CONH-). Peptide bond gives rise to many different types of vibrational modes such as amide A and B bands, amide I, II, III, IV, and VII bands. Among these, amide I and III bands correlated with structural properties of protein molecules yield very prominent Raman bands at 1645~1657 cm^{-1} and 1264~1300 cm^{-1}, respectively, for protein with α-helix structure. The counterpart in protein with β-sheet structure is 1650~1680 cm^{-1} and 1230~1245 cm^{-1}.

Lipids, covering about 12% of total body weight, serve as structural component and energy storage in living organisms. Raman spectroscopy is able to probe biological membrane structure and function without perturbing the sample through probing the major lipids (fatty acids) and its derivatives (phospholipids). They show a lot of vibrational bands in Raman spectra in the region of 100~3000 cm^{-1}, which are structurally sensitive and may be assigned to C-C stretching vibration modes and C-H stretching vibration modes.

Nucleic acids are complex and high-molecular-weight biochemical macromolecules

composed of nucleotide chains that convey genetic information. The most common nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Nucleic acids are found in all living cells and viruses. Raman spectroscopy is extensively used to study the conformation of nucleic acids and mechanism of interaction with other compounds. Nucleic acids show many structurally sensitive Raman bands which can be used to follow the progress of conformational changes or interactions.

Carbohydrates are the most abundant biological molecules and play an important role in living organisms. It stores and transports energy, and also serves as structural components. The vibration modes of carbohydrates are very complex and usually include OH, C-H, C-C, C-C-O, C-O, C-O-C and C-O-H vibrations. This leads the interpretation of Raman spectra to be difficult. For example, C-H stretching vibration of carbohydrates gives a complex pattern in the region of 2800~3050 cm^{-1}. The complexity arises from the presence of different types of CH-containing groups such as $-CH_3$, $-CH_2$, and C-H.

1.2.3.2 Raman Spectroscopic Diagnosis of Human Cancer and Precancer

1. Brain: The early Raman study on human brain tumor tissues was conducted by Mizuno et al with the use of FT Raman spectroscopy in 1994 [6]. Distinctive Raman spectra differences were observed among normal and different types of tumor tissues (i.e., glioma grade II and III, acoustic neurinoma and neurocytoma). Raman band at 960 cm^{-1} due to hydroxyapatite was suggested as a biomarker indicative of tumor pathologies.

Koljenović and co-workers (2002) investigated the feasibility of Raman spectroscopy for grading glioblastoma [7]. They succeeded in delineating cross-sectioned vital and necrotic tissues by using Raman-mapping and K-means cluster analysis (KCA). A

perfect differentiation between necrotic and vital tissues was made by diagnostic model produced by principal component analysis (PCA) and linear discriminant analysis (LDA). In their later work, they demonstrated that HW Raman spectra (2700~3100 cm^{-1}) can provide essentially equivalent diagnostic information as Raman spectra in fingerprint region (400~1900 cm^{-1}) and consequently can be used for tissue characterization [8].

Krafft et al (2005) elucidated the biochemical composition variation among normal and different intracranial tumors (i.e., astrocytoma, glioblastoma multiforme and meningeoma) by modeling tissue spectrum with reference spectra (i.e., protein, lipid, water and cholesterol) [9]. Normal tissue showed higher level of lipid while intracranial tumor tissue showed larger content of hemoglobin and lower ratio of lipid to protein. Recently (2009), they found an increase in water content and a decrease in lipid content in brain tumor as compared to normal brain tissue in a separate study [10].

2. Breast: Redd et al (1993) examined normal and cancerous breast tissues by using Raman spectroscopy with the excitation light at 406.7-, 457.9-, and 514.5-nm [11]. The results showed that Raman signal of normal tissue was mainly attributed to carotenoids (i.e., 1004, 1156, and 1525 cm^{-1}) and lipids (i.e., 1082, 1302, 1444, and 1652 cm^{-1}), which diminished obviously in benign and cancerous breast tissues. Frank and co-workers (1994) optimized the wavelength of excitation light for Raman spectroscopic measurement on human breast tissues [12]. Subsequently, Frank et al (1995) used 784-nm light as the excitation source to study Raman spectral differences between normal and cancerous breast tissues [13].

In the later years, Michael Feld's group did extensive Raman studies on breast cancer.

Manoharan et al (1998) and Chowdary et al (2006) demonstrated that NIR Raman spectroscopy coupled with multivariate statistical techniques was capable of differentiating breast cancer from normal tissue with a fairly good accuracy [14, 15]. To further elucidate the biochemical composition changes associated with tissue malignant transformation, Shafer-Peltier et al (2002) developed a biochemical and morphological model of breast tissue with the component spectra derived from spectra of cell cytoplasm, cell nucleus, fat, β-carotene, collagen, calcium hydroxyapatite, calcium oxalate dehydrate, cholesterol-like lipid deposits and water [16]. Haka et al (2005) extended this study on a larger women population (58 patients) and achieved a sensitivity of 94% and specificity of 96% in discriminating cancer tissues from normal, fibrocystic and fibroadenoma tissues by performing logistic regression analysis on the fitting coefficients [17]. In 2006, Haka et al reported an *in vivo* study on using Raman spectroscopy to delineate the malignancy margin during partial mastectomy and their success in real-time Raman spectroscopic diagnosis [18]. Three years later, with the use of the same system, they did the first trial on the prospective diagnosis of breast cancer using Raman spectroscopy [19]. In the same year (2009), Chowdary et al employed nonlinear peak fitting, known as Lavenberg-Marquardt method, to deconvolve tissue spectrum into 17 individual Raman bands to elucidate the biochemical changes induced by tissue cancerous changes [20].

Besides cancer detection, calcification also received some research efforts. Haka et al (2002) investigated the feasibility of identifying micro-calcification in benign and malignant breast lesions using a NIR Raman microscope [21]. A sensitivity of 88% and specificity of 87% were achieved in differentiating micro-calcification between benign and malignant breast tissues. Matousek et al (2007) demonstrated that transmission Raman spectroscopy has the potential to aid conventional screening methods (e.g.,

mammography and ultrasound) in improving the early diagnosis of breast cancer by probing the calcification in breast tissues [22].

3. Colon and Esophagus: Keller and co-workers (1994) measured Raman spectra from a pair of normal and haemorrhagic intestines using FT Raman spectroscopy with 1064-nm light excitation [23]. Haemorrhage-induced Raman spectral differences were observed between these two tissue types. Feld et al (1995) reported another preliminary study on Raman spectroscopic examination of normal and cancer colon tissues [24]. Obvious spectral difference between normal and adenocarcinoma colonic tissues occurred near 1000, 1300 and 1500 cm^{-1}. Shim et al (2000) succeeded in acquiring Raman spectrum from gastro intestine (GI) *in vivo* for the first time using a flexible fiber-optic Raman probe through the conventional endoscope [25]. Prominent Raman peaks were seen in the vicinities of 1003 (phenyl ring breathing mode), 1260 (amid III), 1310 (CH_2 twisting mode), 1450 (CH_2 bending mode), and 1657 cm^{-1} (amide I). By using the same Raman system, Molckovsky (2003) demonstrated the capacity of Raman spectroscopy for identifying hyperplastic and adenomatous polyps both *ex vivo* and *in vivo* [26]. Andrade et al (2007) demonstrated the existence of intrinsic spectral variation and suggested that this variation should be taken into consideration when building spectral database [27]. Widjaja et al (2008) investigated the multi-type diagnostic ability of NIR Raman spectroscopy for colon tissues on a large dataset [28]. Diagnostic accuracy of 99.9% was yielded by a diagnostic algorithm based on PCA and support vector machine (SVM).

Shim et al (2000) demonstrated the feasibility of measuring Raman spectra from esophagus tissues *in vivo* with the use of fiber-optic probe via conventional endoscope [25]. Then, Song et al (2005) proved the *in vivo* diagnostic potential of NIR Raman spectroscopy on a large population of 65 patients, showing an accuracy of about 85%

in diagnosing dysplasia [29]. Kendall et al (2003) employed a NIR micro-Raman system to test if Raman spectroscopy can identify neoplasia in Barrett's esophagus [30]. Shetty and co-workers (2006) further elucidated the biochemical changes in carcinogenesis of esophagus through Raman spectroscopic mapping [31]. The major finding was the reduced content of glycogen and elevated content of DNA in abnormal area.

Recently, micro-Raman probe was proposed for acquiring the Raman signal originating from the subsurface of esophagus by Hattori et al (2007) [32]. Day and co-workers (2009) optimized a confocal Raman probe design for *in vivo* Raman measurement on esophagus via the auxiliary channel of conventional endoscope [33]. The probe has been validated *ex vivo*, showing the ability of acquiring Raman spectrum from esophagus with 2 seconds.

4. Bladder and Prostate: Nie and co-workers (1992) demonstrated for the first time that Raman spectroscopy can be used to discriminate bladder cancer from normal tissue [34]. Raman spectra of rat bladder tissues measured with NIR FT Raman spectroscopy exhibited prominent bands at 1004 (phenylalanine), 1240 and 1268 (amide III), 1449 (lipid) and 1664 cm^{-1} (amide I). Jong et al (2002) proved that Raman spectroscopy in combination with cluster analysis can characterize bladder wall layers, including urothelium, muscle and lamina [35]. Subsequently, they (2003) extended their study to investigate the effect of outlet obstruction on the molecular composition of bladder muscle tissue by using Raman spectroscopy [36]. They observed collagen infiltration and an accumulation of glycogen in obstructed bladder tissues. Crow et al conducted a lot of work on applying Raman spectroscopy for the detection of both bladder and prostate cancer and elucidating the biochemical changes, such as multi-classification among benign, and three grades of prostatic adenocarcinoma

tissues [37] and among normal bladder, cystitis and carcinoma in situ (CIS)/translational cell carcinoma (TCC) [38]. Later, a fiber-optic Raman system was developed for *in vivo* Raman measurement on bladder and prostatic tissue and has been evaluated *in vitro* [39]. In addition to the study on gross tissues, Crow and co-workers (2005) investigated the feasibility of Raman spectroscopy coupled with PCA-LDA to characterize various prostatic cancer cell lines [40]. Shortly after their study, Taleb et al (2006) employed partial least square discriminant analysis and adjacent band ratios to successfully classify Raman spectra of all malignant prostatic cells from that of normal cells [41].

Jong and co-workers (2006) tried to delineate the normal and tumor part of bladder tissues with PCA and hierarchical cluster analysis (HCA) [42]. Furthermore, the biochemical difference between non-tumor and tumor tissues was elucidated by modeling tissue spectra with 18 reference spectra of biochemical representatives. Stone et al (2007) used similar chemical fitting method to study the biochemical composition variation related to malignant transformation of bladder and prostatic tissues with 8 reference spectra of actin, collagen I/III, choline, triolein, oleic acid, DNA and cholesterol [43].

5. Larynx and Nasopharynx: An early study on Raman spectroscopic diagnosis of malignant changes in larynx was conducted by Stone and co-workers (2000) [44]. Seven normal, four dysplastic and four malignant larynx tissues were examined using Raman microscope with 830-nm light and discriminated with an accuracy of more than 85% by PCA-LDA diagnostic algorithm. Five years later, in another study, Lau et al (2005) demonstrated again that Raman spectroscopy can be used for identifying larynx tissues [45]. Teh et al (2009) proposed random forest method to develop diagnostic algorithm for laryngeal cancer [46]. Random forest method can also provide insight

into the biochemical changes associated with tissue malignant transformation.

As for nasopharynx, two studies were reported in literature. Lau et al (2003) firstly reported their preliminary findings on Raman spectroscopic characterization of nasopharynx cancer. In recent year, Feng et al (2009) studied the Raman features of cancerous and normal nasopharynx tissues by using surface-enhanced Raman scattering (SERS) technique [47]. The Raman images created with the Raman bands at 725, 962 and 1336 cm^{-1} showed an increased, reduced and increased intensities, respectively, in cancerous nasopharyngeal tissues as compared to normal tissues.

6. Lung: Raman spectroscopy was firstly used to study lung tissue changes other than cancer, such as inclusions identification by Buiteyeld et al (1984) [48]. In a later study, Schut and co-workers (1997) demonstrated that Raman spectroscopy can probe the change of carotenoid content level in lung tissue [49]. The first direct comparison of Raman spectra between normal and cancerous lung tissues was performed by Kaminaka et al (2001) and showed an increased intensity of Raman peaks at 1448 and 1666 cm^{-1} due to collagen in cancer tissue [50].

Huang et al (2003) examined bronchial tissues (12 normal, 10 squamous cell carcinoma (SCC) and 6 adenocarcinoma) from 10 patients using Raman spectroscopy [51]. Raman peak intensity ratio of 1445/1655 cm^{-1} (CH_2 scissoring/collagen) was found to be an effective diagnostic marker with a sensitivity of 92% and specificity of 94%. In the same year, they (2003) reported that formalin-fix process had effect on tissue Raman spectra and hence suggested the use of fresh tissue for Raman study [52]. Following Huang's *ex vivo* work, Short et al (2008) successfully designed a flexible fiber-optic endoscope Raman probe for *in vivo* Raman measurement on bronchial tissue [53].

In addition to Raman spectra classification, Koljenović and co-workers (2004) made efforts to gain the understanding of biochemical and morphological composition of bronchial tissue using Raman microscopy [54]. Raman map of cross-sectioned bronchial tissues was created by implementing PCA and KCA on Raman spectra. The Raman map showed a good agreement with the corresponding histology. Krafft et al (2008) used similar method to make a pair-wise comparison of Fourier transform infrared spectroscopy (FTIR) and Raman imaging between normal and congenital cystic adenomatoid malformation (CCAM) lung tissues [55]. It was found that CCAM tissues contained an increased lipid and glycogen content and reduced red blood cell content as compared to normal tissues.

The recent studies on Raman spectroscopic diagnosis of lung precancer and cancer were performed at cellular levels. Jess and co-workers (2009) measured Raman spectra from normal, neoplastic lung cell lines (mild and severe neoplasia) successfully using Raman microspectroscopy [56]. Neoplastic cell lines were discriminated from normal cell lines with a sensitivity of 91% and specificity of 75%. Moreover, neoplastic cell lines were further graded into two stages (i.e., mild and severe) with an accuracy of 79% and 87%, respectively. Similar study on separating lung cancer cells from normal cells with Raman spectroscopy was reported by Oshima et al (2010) [57]. Eighty percent of the cell lines were characterized correctly with PCA.

7. Skin: Basal cell carcinoma (BCC) is the most common form of skin cancer and therefore has received most of Raman research efforts. Gniadecka and co-workers (1997) carried out Raman measurements on both normal and BCC tissues by using NIR FT Raman spectroscopy [58]. Raman spectral features in the regions of 830~900 cm^{-1}, 900~990 cm^{-1}, and 1220~1300 cm^{-1} allowed a complete separation between BCC and normal tissues, which was confirmed by neural networks analysis. Nijssen et al

(2002) obtained a sensitivity of 100% and specificity of 93% in discriminating BCC from its surrounding non-cancerous tissues using PCA and logistic regression analysis [59]. In their following work, they (2007) demonstrated that NIR Raman spectroscopy in HW region (2800~3125 cm^{-1}) was also capable of characterizing BCC [60]. Tissue spectral modeling with reference spectra of collagen, oleic acid, palmitic acid and albumin revealed a reduced collagen content and increased albumin content in BCC tissues as compared to normal tissues. With the similar method, Short et al (2006) found that the nucleoli from tumor cells contained less RNA, histone, and actin than that from normal cells while more DNA, histone, and actin for the remaining nucleus [61]. Ly and co-workers (2008) demonstrated that polarized Raman spectroscopy can improve the differentiation between normal and peritumoral dermis as compared to conventional non-polarized Raman spectroscopy [62]. In a recent study, Larraona-Puy et al (2009) used Raman spectroscopy not only to distinguish BCC from normal tissues but also to delineate the tumor margin [63].

Besides BCC, Raman spectroscopy has also been applied to detect other forms of skin cancer, especially melanoma which is the most aggressive skin cancer. At the early time, Gniadecka and co-workers (1997, 2004) explored the Raman spectral variation among a big variety of skin pathologies and employed neural networks technique to distinguish the Raman spectra of different tissue types [64, 65]. Melanoma was separated from pigmented nevi, BCC, seborrheic keratoses and normal skin with a sensitivity of 85% and specificity of 99%. A perfect prediction of normal, BCC, SCC and melanoma by Raman spectroscopy was reported by Lieber et al (2008) [66].

Meanwhile, several *in vivo* Raman studies on skin have been reported. Huang et al (2001) developed a rapid NIR Raman spectroscopy coupled with a fiber-optic probe for *in vivo* Raman measurement on skin and acquired good-quality cutaneous Raman

spectra within 1 second [67]. More recently, Lieber et al (2007) developed a handheld *in vivo* Raman microspectrometer [68]. In a subsequent study, this system was used for clinical measurement by the same group [69]. Maximum representation and discrimination feature were used to reduce data dimension and sparse multinomial logistic regression was used to develop diagnostic algorithm. All abnormal skin tissues were predicted correctly and only two normal tissues were misdiagnosed.

8. Stomach: Teh and co-workers have conducted extensive studies on applying Raman spectroscopy for the *ex vivo* detection of precancer and cancer in stomach from 2008 till now. Their first report (2008) was made on differentiating dysplasia from normal tissue using NIR Raman spectroscopy [70]. PCA-LDA analysis yielded a sensitivity and specificity of 95.2% and 90.9%, respectively. In 2009, they extensively investigated the potential of empirical method (i.e., intensity ratio) for distinguishing stomach dysplasia [71]. The combination of I_{875}/I_{1450} and I_{1208}/I_{1655} proved optimal for the diagnosis of dysplasia with a sensitivity of 90.5% and specificity of 90.9%.

They (2008) also attempted to identify stomach cancer using Raman spectroscopy and classification and regression tree (CART) technique instead of PCA-LDA [72]. The CART algorithm yielded a sensitivity of 90.2% and specificity of 95.7%, and moreover found that Raman bands at 875 and 1745 cm^{-1} were the most significant Raman features for cancer discrimination. In their latest work, they (2010) tested the possibility of further typing stomach cancer (i.e., intestinal and diffuse adenocarcinomas) using Raman spectroscopy [73]. The correct prediction rates are 88%, 92% and 94% for normal, intestinal type and diffuse type tissues, respectively. In addition to precancer and cancer, they (2010) also did a pilot study to characterize nonneoplastic stomach lesions (i.e., *Helicobacter-pylori* (*Hp*) infection and intestinal metaplasia (IM)), which are highly associated with stomach cancer [74]. Good

differentiation among normal, IM and *Hp*-infection tissues was achieved, showing accuracies of 91.7%, 80.0%, and 80.0%, respectively. Besides Teh and co-workers, several other groups were also involved in applying Raman spectroscopy for stomach cancer diagnosis [75-77].

Recently, Huang et al (2009) reported their success in measuring Raman spectrum from stomach *in vivo* [78]. A fiber-optic endoscopic Raman probe was designed, which was flexible and compatible with conventional endoscopy, and consequently allowed *in vivo* Raman measurement on stomach. Moreover, white, autofluorescence and narrow band imaging was used to guide the Raman measurement *in vivo*.

1.3 Cervical Cancer

1.3.1 Cervical Cancer Facts and Risk Factors

1.3.1.1 Cervical Cancer Facts

Cervical cancer is the 2nd most frequent cancer among women worldwide in 2002 and shows an incident rate of 16%. It is ranked 7th, 2nd and 4th in developed countries, developing countries and Singapore, respectively [1]. The corresponding incidence rate is 13.6%, 16.6% and 15.7%, respectively. More importantly, the incidence rate is still growing [1]. Estimated based on the rate in 2002, the incidence rate in 2010 is 17.2%, 14.1%, 18.3% and 19.4% worldwide, in developed and developing countries, and Singapore, respectively. The number of new cervical cases in 2010 is 585,278, 88,702, 505,592 and 441, respectively, while just 493,243, 83,437, 409,404 and 323 in 2002, respectively [1].

Meanwhile, the mortality rate of cervical cancer is in the 3rd place following breast and lung worldwide. In particular for developing countries, it is the highest [1]. In

Singapore, it is the 5th highest. In 2002 and 2010, the specific mortality rate of cervical cancer is 8.9% and 9.7% in the world, 6.4% and 6.9% in developed countries, 9.5% and 10.5% in developing countries, and 9.9% and 13.3% in Singapore, respectively. Simultaneously, the number of deaths due to cervical cancer grows from 273,505 to 327,899 in the world, 39,512 to 43,043 in the developed countries, 233,776 to 291,872 in developing countries, and 205 to 302 in Singapore [1]. Even given that the incidence and mortality rates remain unchanged, the absolute number of new cancer cases and cancer-induced deaths still keeps increasing. This may be accounted for by the increased population size and population aging [79]. Therefore, prevention and early diagnosis of precancer and cancer in the cervix are becoming even more desired, especially in developing countries where 80% of the cervical cancers occurs [80].

1.3.1.2 Risk Factors

A variety of factors have been found to separately or jointly lead to cervical cancer, such as human papillomavirus (HPV) infection, smoking, oral contraceptive, the number of sex partners, the number of full pregnancies, genetic and immunological factors. Among the factors above, HPV infection has been recognized as the main cause of cervical cancer during the past twenty five years. The eight most common HPVs (i.e., HPV-16, -18, -33, -45, -31, -58, -52 and -35 in order of decreasing prevalence) account for ~90% of cervical cancer worldwide and HPV-16 and -18 account for 70% of cervical cancer [81]. Consequently, HPA DNA testing and HPV vaccine have emerged as the choice of method for preventing cervical cancer.

Smoking is another major factor of cervical cancer, leading to a 2-fold increase in the risk of cervical cancer [82]. Moreover, the risk also shows a trend of increase with the increase in the number of cigarette smoked and years of smoking [82]. Besides, a long

term use of oral contraceptive (i.e., >5 years) and the number of full pregnancies (i.e., >=7) induce a significant increase up to 4-fold in the risk of cervical cancer of HPV-infected patients [83, 84]. Genetic predisposition, the host response and immune-suppression are also found to contribute in part to the development of precancer and cancer [85, 86]. Diet, education and social class are found not to be significantly associated with cervical precancer and cancer [87, 88].

1.3.2 Anatomy of Cervix

The cervix is the lower narrow part of the uterus and protrudes through the upper wall of the vagina. It serves to transport the menstrual blood from uterus to vagina and direct sperm into uterus during intercourse. The cervix is further divided into ectocervix and endocervix. Ectocervix is the portion of the cervix which projects into the vagina. It is covered by stratified non-keratinizing squamous epithelium which is centered in the external orifice of the uterus (os). The ectocervix is around 3 cm long and 2.5 cm wide on average and usually visible under colposcopy [89]. In comparison, the endocervix is the part of the cervix connecting external os and uterine cavity. It is covered by columnar epithelium. The joint point between ectocervix and endocervix is called squamo-columnar junction. The portion of the columnar epithelium that is ultimately replaced by squamous epithelium is termed the transformation zone, where neoplasia and malignancy arise from [89].

The cervix shows a two-layer structure, comprising the superficial epithelium layer and the underlying stroma layer. The stroma layer is composed predominantly of elastic tissue (i.e., collagen and elastin) forming extracellular matrix. The epithelium is separated from the stroma by basement membrane. It differs between ectocervix and endocervix. The epithelium of ectocervix shows architecture of multiple cell layers,

including superficial layer, intermediate layer, parabasal layer and basal layer from the epithelium surface towards the stroma. In contrast, the endocervical canal is covered by columnar epithelium, which is composed of a single layer of tall cells.

1.3.3 Histology of Cervix

1.3.3.1 Normal Cervix

The squamous epithelium of normal cervix shows a clear multi-layer structure, consisting of basal layer, parabasal layer, intermediate layer, superficial layer [90]. The basal layer consists of cells in cylindrical shape with relatively large nuclei. The parasal layer is formed by a couple of cell layers with fairly large nuclei and distinct intercellular bridges. Mitotic figures are rarely seen in these two layers. The intermediate layer on top of parabasal layer is formed by flatten cells with a high level of glycogen and frequent vacuolation in cytoplasm. The superficial layer is composed of elongated, flattened cells with small pyknotic nucleolus and a large amount of cytoplasm.

1.3.3.2 Cervical Dysplasia

Cervical dysplasia refers to the cervical intraepithelial changes which show malignant potential. It was termed as cervical intraepithelial neoplasia (CIN) by Richard in 1967 [91], and is the premalignant stage of cervical cancer. Histological features of CIN mainly include the absence of cytoplasmic differentiation and orderly stratification, the lack of clearly defined boundaries, and large nuclei-cytoplasm ratio in epithelium cells. CIN can be graded into three stages (i.e., 1, 2 and 3) according to the spread of neoplastic changes in the epithelium. CIN 1 shows neoplastic changes within the lower one-third of the epithelium while lower two-thirds and whole epithelium for CIN 2 and 3, respectively. Besides, CIN 3 also shows undifferentiated, non-stratified, basaloid

cells with nuclear crowding and greater nuclear pleomorphism as compared to CIN 1 and 2. However, CIN does not show any invasion into the underlying stroma, which serves as an important criterion for distinguishing CIN from invasive carcinoma.

Detection of CIN at mild stage followed by effective treatment can reduce incidence rate of cervical cancer. However, the diagnostic inconsistency, especially on CIN 1 and 2 still exists [92]. Therefore, a modified two-tier system (the Bethesda system [93]) was proposed, which divided cervical epithelial neoplastic changes into two groups including low grade and high grade squamous intraepithelial lesions (LGSILs and HGSILs). LGSILs refer to HPV infection and CIN 1, and HGSILs refer to CIN 2 and 3, and carcinoma in situ (CIS). However, the discrimination between CIN 1 and 2 is not improved under the two-tier system. Consequently, a robust and objective diagnostic method is desirable to prevent the aforementioned diagnostic inconsistency among pathologists. Optical methods have proven to be a potential candidate.

1.3.3.3 Cervical Cancer

Cervical cancer can be usually differentiated from precancer by the invasion of the neoplastic changes into the stroma. Cervical cancer mainly has two forms. One is squamous carcinoma, constituting 70~78% of cervical cancer. Squamous carcinoma shows some of the following characteristics, including relatively large cells, bands, discrete islands, infiltrative pattern and/or solid sheets [94]. The other is adenocarcinoma, constituting 12~18.6% of cervical cancer [94]. Little or no cytologic atypia is observed in adenocarcinoma and the cells are like that of normal counterpart. Glandular crypts are often sharply angulated and extend towards a big depth of the stroma. In addition, the surrounding stroma shows a loose edematous or desmoplastic response.

1.3.4 Conventional Screening/Diagnosis and Treatment of Cervical Cancer

Cervical cancer has a defined premalignant phase for many years. This allows a long time for screening or diagnosis of premalignant changes or malignancy at early stage. Then, cervical cancer can be prevented and the corresponding mortality rate can be reduced significantly. Therefore, various cervical cancer screening methods are developed. To date, cervical cytology testing (i.e., pap smear) is widely used as an effective screening method for cervical cancer, especially in developed countries. The procedures of pap smear include collecting cells from cervix and then examining those collected cells under microscopy. However, the use of pap smear in developing countries is limited and shows a deteriorated efficacy. Consequently, visual inspection techniques, such as visual inspection after applying acetic acid (VIA), visual inspection after applying Lugol's iodine (VILI), are proposed for routine cervical cancer screening method as an alternative to pap smear in the developing countries. Recently, HPV DNA testing is emerging as another alternative method due to the high correlation of HPV with cervical cancer. Moreover, HPV DNA testing can be done on site and yield the results immediately. As the advance of HPV DNA testing assay, it is growing as a cost-effective, efficient and suitable cervical cancer screening technique, especially in developing countries. More recently, HPV vaccine is developed for preventing cervical cancer. However, most of the HPV vaccines can only be targeted at limited types of HPVs.

For the diagnosis of cervical precancer and cancer, colposcopic examination and colposcopy-directed biopsy remain as the most reliable method. Colposcopy is like a telescope which can magnify the cervix for examination. In clinic procedures, 3~5% acetic acid will be applied on the cervix for a few minutes prior to checking the colposcopic appearance of cervix. Suspicious lesions found under colposcopy will be

biopsied for histopathological examination. Histopathology results serve as the gold standard for cervical precancer or cancer diagnosis.

At present, almost all the precancers in the cervix are curable. The treatment strategy varies with the severity and size of the lesion. The common treatment methods for cervical precancer include loop electrosurgical excision procedure, large loop excision of the transformation zone, hysterectomy, laser therapy (vaporizing cervical epithelium with laser) and cryotherapy [95, 96].

The treatment of invasive cancer depends on the stage of invasive cancer. For stage I, radical hysterectomy is preferred and minimally invasive surgery recently emerges as an alternative for small tumors. For stage II to IV, radiotherapy is usually used. In addition, chemotherapy is proposed as an adjunct to radiotherapy to improve progression-free survival [95].

1.4 Raman Spectroscopic Diagnosis of Cervical Cancer

Liu and co-workers pioneered Raman measurement on cervical tissues in 1992, including malignant and non-malignant (i.e., normal and benign) tissues with NIR FT Raman spectroscopy [97]. Their most significant finding was that Raman peak at 1657 cm^{-1} (amide I) showed a higher intensity than Raman peak at 1445 cm^{-1} (C-H bending of proteins/lipids) in normal tissue while it showed otherwise in malignant tissue. Mahadevan-Jansen et al (1998) demonstrated the capacity of Raman spectroscopy for the detection of cervical precancer using both intensity ratio method and multivariate statistic technique (i.e., PCA and fisher discriminant analysis (FDA)) [98]. The Raman peak intensity and intensity-ratio, including I_{1070}, I_{1656}, I_{1656}/I_{1330} and I_{1656}/I_{1454} were found to be diagnostically significant for differentiating squamous intraepithelial lesions (SILs) from non-SILs as well as HGSILs from LGSILs. In their following

work, they proceeded to the *in vivo* Raman spectroscopic diagnosis of cervical dysplasia with the design of a fiber-optic Raman probe [99]. Raman spectra with acceptable signal to noise ratio (SNR) can be acquired from cervix *in vivo* within 90 seconds. Subsequently, clinical trials were conducted on 13 patients, including normal, inflammation, squamous metaplasia, LGSILs and HGSILs [100]. A diagnostic algorithm of I_{1454}/I_{1656} versus I_{1330}/I_{1454} yielded a good differentiation of HGSILs from other tissues.

Robichaux-Viehoever (2007) employed an improved Raman system for *in vivo* Raman measurement on cervix with integration time reduced to a few seconds [101]. Logistic regression analysis was employed to utilize the intensity of the primary peaks (i.e., 1006, 1055, 1244, 1305, 1324, 1450, 1550 and 1657 cm^{-1}) to develop diagnostic algorithm, showing a better performance than colposcopic diagnosis. In a later report, Kanter et al (2009) developed a novel statistical method for multi-group classification [102]. Maximum representation and discrimination feature were used to extract diagnostic information and sparse multinomial logistic regression was used to make differentiation. With this method, fairly high correct prediction rate was achieved for HGSILs (24/29), LGSILs (18/21), metaplasia (19/29) and normal (208/226). Using the same technique, they (2009) investigated the effect of hormonal variation on Raman spectra for tissue characterization [103]. The results showed that the accuracy of discriminating LGSILs from normal ectocervix could be improved by stratifying the tissues according to the menopausal status.

Besides Anita's group, several other groups also contributed to applying Raman spectroscopy for the detection of cervical precancer and cancer. Jess et al (2007) demonstrated that Raman spectra of various cervical cell lines (i.e., normal and malignant cell lines, and cell lines expressing HPV16 E7) can be separated well by

PCA-based algorithm [104]. They also found that cell fixation resulted in a large reduction in peak intensities at 936 and 1090 cm^{-1} and moreover can increase the sensitivity of cell characterization compared to live cell lines. Martinho et al (2008) demonstrated that cervicitis may deteriorate the performance of Raman spectroscopy for the diagnosis of cervical precancer [105].

For cervical cancer diagnosis, Krishna and co-workers (2006-2007) explored the potential of Raman spectroscopy for cervical cancer detection [106, 107]. The major spectral variation associated with tissue malignancy included narrowed amide I and III bands, a minor red shift of Raman band at 1450 cm^{-1}, and the vanishing of Raman bands at 854, 939, 1269 and 1384 cm^{-1}. Diagnosis with the inclusion of the parameters (factor scores, Mahalanobis distance and spectral residual) produced by PCA achieved a very high sensitivity and specificity of 99.5% in distinguishing cancer from normal tissues. Moreover, they also proved that Raman spectroscopy can be used to evaluate the efficacy of radiotherapy [107]

Lyng et al (2007) elucidated the biochemical changes associated with tissue malignant transformation with reference spectra of proteins, nucleic acids, lipids and carbohydrates [108]. An absence of glycogen band and an increased intensity of nucleic acids and amide I bands were observed in cancer tissue as compared to normal tissue. About 99% of normal, dysplasia and cancer tissues can be characterized correctly by PCA-LDA diagnostic algorithm. A similar study was reported by Kamemoto et al (2010) [109]. They observed that normal squamous cells exhibited a more intense Raman scattering than malignant squamous cells. The spectra in the region of 775~975 cm^{-1} exhibited a good correlation with the spectra in the region of 2800~3100 cm^{-1}.

1.5 Other Optical Spectroscopic Techniques for Cervical Cancer Diagnosis

1.5.1 Fluorescence Spectroscopy

Lohmann and co-workers (1989) measured fluorescence spectra from normal, dysplasia, and invasive carcinoma cervical tissues with 365-nm excitation light [110]. Tissue fluorescence spectra showed a primary band peaked at 475 nm, which was attributed to reduced form of nicotinamide adenine dinucleotide (NADH). The intensity of NADH band increased with the progression of tissue from normal to severe dysplasia and was very low in tumor tissue. This was confirmed by the fluorescence image of cervical tissue cryo-section [111]. The feasibility of using fluorescence spectroscopy for discriminating cervical malignancy was explored for the first time by Glassman et al (1992, 1994) [112, 113]. The results indicated that the excitation band ratio (I_{335}/I_{380}) and emission band ratio (I_{340}/I_{440}) can be a biomarker of cervical cancer. Mahadevan et al (1993) recorded fluorescence excitation-emission matrices (EEMs) *in vitro* on 18 biopsied tissues of 10 patients [114]. Fluorescence intensity at 380 nm under 330-nm light excitation can distinguish dysplasia with a comparable accuracy as compared to colposcopic diagnosis. Koumantakis et al (1997) found other fluorescence spectral bands at various wavelengths (i.e., 558-, 583-, 600-, 630- and 697-nm) under 420-nm excitation as cervical malignancy predictors [115].

In the following years, Ramanujam and co-worker conducted an extensive work on applying fluorescence spectroscopy for the *in vivo* detection of cervical dysplasia with 337-nm excitation light [116-119]. In their first *in vivo* clinical trial, fluorescence spectra were acquired from the cervical tissues of 28 patients, including 66 colposcopically normal and 49 histologically abnormal (i.e., inflammation, HPV

infection, and CIN) using a bifurcate fiber probe [116, 117]. Dysplasia tissue can be distinguished with an accuracy of ~90% using overall fluorescence intensity and the slope of the spectra in the region of 420~440 nm. Biochemical composition changes were elucidated by fitting tissue spectra with reference spectra of pure biochemicals [117]. Reduced collagen content occurred in dysplasia while oxyhemoglobin attenuation and NADH content were increased. In a later study, they developed multivariate statistical algorithm for the diagnosis of cervical dysplasia using PCA and logistic discriminant analysis, showing a sensitivity of 88%±1.4 and specificity of 70%±1 [118]. This multivariate statistic method was further used to evaluate the performance of 380- and 460-nm light excited fluorescence in diagnosing cervical precancer [119]. Fluorescence spectroscopic diagnosis with 460-nm excitation outperformed over that with 380-nm excitation.

Subsequently, several groups were involved in investigating the various variables associated with i*n vivo* clinical trial on fluorescence spectroscopic diagnosis of cervical precancer. Utzinger et al (1999) studied whether tissue type, sample size, population and SNR affected the fluorescence spectroscopic diagnosis of cervical precancer [120]. In the same year, Agrawal et al (1999) investigated the effect of acetic acid, cervical mucus and vaginal medication [121]. Chang et al (2002) found that menstrual cycle did not result in a significant inter-patient variation and hence menstrual cycle did not need to be considered during fluorescence measurement [122]. However, Cox et al (2003) observed a reduction in the fluorescence emission in the first several days of one cycle and suggested no fluorescence measurement during those days [123]. Gill et al (2003) pointed out that menopausal status had a significant effect on the fluorescence spectra of cervical tissues, such as a significantly higher stroma fluorescence for postmenopausal women [124]. Atkinson et al (2005) found that

overall fluorescence spectral intensity increased with the increase of patient age and Follicle Stimulating Hormone [125]. Meanwhile, Brookner et al thought that fluorescence spectral difference between the normal areas of cervix of patients with and without dysplasia history was statistically significant [126]. Nath et al (2004) and Rivoire et al (2004) reported that fiber probe pressure did not cause significant effect on the fluorescence intensity and line-shape [127, 128]. Freeberg et al (2007) described the device variability of fluorescence and reflectance spectroscopy and the induced effect on the spectral measurement [129]. The spectral data analysis should be performed on the stratified data as the spectral intensity varied with devices. This is in agreement with the findings reported by Pikkula et al (2007) [130].

Fluorescence imaging has been widely used to study tissue fluorescence properties associated with malignancy. Brookner et al (2000) observed that fluorescence intensity decreased in epithelium but increased in stroma of normal cervix with the increase of patient age [131]. Subsequently, fluorescence image pattern was correlated with tissue premalignancy and malignancy. Ramanujam et al (2001) recorded the fluorescence image of cross-sectioned cervical tissues with 365- and 440-nm light excitation [132]. The severe dysplasia exhibited the lowest intensity of the epithelium fluorescence, intensity ratio of epithelium/stroma and redox ratio. Drezek and co-workers (2001) reported an increase in NADH fluorescence and a reduction in collagen fluorescence with tissue malignant transformation [133, 134]. Similar studies and findings were also reported by other groups [135-138]. Besides NADH and collagen, Chang et al (2006) found that flavin adenine dinucleotide (FAD) fluorescence increased and keratin fluorescence decreased in the epithelium of dysplasia tissue relative to normal tissue through quantitative analysis of tissue fluorescence [139].

Chang et al (2002) optimized the excitation light wavelength (i.e., 330~340, 350~380

and 400~450 nm) for fluorescence spectroscopic diagnosis of cervical precancer [140]. In the following year (2003), Benavides et al reported their development of a multispectral digital colposcopy (MDC) for *in vivo* detection of cervical cancer [141]. Dysplasia area can be identified with automated diagnostic algorithm. An extensive use of MDC was conducted on 46 patients by Milbourne et al (2005) [142]. The MDC images can be matched to histopathologic and algorithmic (non-parametric K-nearest neighbor classifier) maps, indicating the diagnostic potential of MDC.

Apart from the literature above, there are some other works reported on fluorescence spectroscopic diagnosis of cervical precancer and cancer. Heintzelman et al (2000) attempted to explore the fluorescence difference between inflammation and cancer in cervix by using EEMs [143]. Nordstrom et al (2001) achieved a differentiation of SILs from normal squamous tissues with an accuracy of higher than 85% [144]. Weingandt et al (2002) reported that fluorescence intensity under 375- and 440-nm excitation decreased with tissue premalignant transformation [145]. Rodero et al (2008) proposed a simple intensity-based algorithm to differentiate fluorescence spectra under 488-nm light excitation from chronic inflammation, LGSILs and HGSILs [146]. Chidananda et al (2006) achieved an accuracy of higher than 95% in differentiating cervical malignancy from normal tissue using fluorescence spectroscopy and PCA [147]. The ratio of collagen/NADH showed a potential in predicting malignant tissue.

1.5.2 Reflectance Spectroscopy

Much less work has been done to use reflectance spectroscopy for cervical precancer and cancer detection as compared to fluorescence spectroscopy. Nordstrom et al (2001) attempted to characterize cervical precancer and benign changes using diffuse reflectance spectroscopy coupled with Mahalanobis distance [144]. Reflectance

spectroscopy showed an inferior diagnostic performance to fluorescence spectroscopy. Georgakoudi and co-workers (2002) measured *in vivo* reflectance spectra on cervix and extracted the scattering coefficient using a diffuse-model [148]. A sensitivity of 62% and specificity of 82% in discriminating SILs from non-SILs was produced by logistic regression algorithm calculated from the slope and intercept of the reduced scattering coefficient function curve. Moreover, the diagnosis was improved by combining reflectance with fluorescence and light scattering spectra. Chang et al (2005) agreed on the superiority of fluorescence spectroscopy to reflectance; however, they did not see any improvement in the diagnosis by combining reflectance and fluorescence [149]. Mirabal et al (2002) combined the reflectance spectra of cervix measured at four different illumination-collection separations as the input of statistical analysis (i.e., PCA) [150]. The differentiation of normal squamous from HGSILs and normal columnar from HGSILs was found to be comparable to that based on fluorescence spectra. Marín et al (2005) extended Mirabal's work to further explore the reflectance spectral pattern related to precancer on a larger patient population [151]. Precancer tissues showed weaker reflectance intensity as compared to normal tissues. They also suggested that Soret absorption of hemoglobin and the wavelength-dependent slope of spectra should be the most significant diagnostic markers.

To better understand the *in vivo* reflectance spectra of cervix, Arifler et al (2006) proposed MC simulation method to model the *in vivo* spectra [152]. The effect of optical properties (i.e., scattering and absorption coefficients) on the reflectance spectra and the depth-selectivity of the probe were revealed. In a later report, Weber et al (2008) described an adjoint reflectance and fluorescence model to fit the *in vivo* measured spectra of cervix [153]. Volume fraction of blood, hemoglobin oxygen

saturation, concentration of structural protein, the strength and shape of the scattering in the stroma represented almost all the diagnostic information in reflectance spectra. Chang and co-workers (2009) also used MC method to extract the physiological parameters from diffuse reflectance spectra [154]. It was found that total hemoglobin concentration was significantly higher in HGSILs than in LGSILs and normal while reduced scattering coefficient was reduced in SILs compared to normal tissues.

1.5.3 Infrared Spectroscopy

Wong and co-workers (1991) pioneered the use of IR spectroscopy for investigating the structural changes of dysplastic and malignant cervical cells and tissues [155, 156]. The significant molecular structural changes associated with tissue malignant transformation included: (1) reduction in glycogen; (2) extensive hydrogen bonding of the phosphodiester groups of nucleic acids; (3) red-shift of IR band at 1082 cm^{-1}; (4) reduction in hydrogen bonding of C-OH groups of carbohydrates and proteins; (5) increased degree of disorder of methylene chains of membrane lipids; (6) an additional band at 970 cm^{-1}; (7) decreased methyl-to-methylene ratio; (8) hydrogen-bond strength amide groups decreased in α-helical segments but increased in β-sheet segments. Wood et al (1996) employed PCA to discriminate the IR spectra of exfoliated cells from 272 patients [157]. Eighty-six percent of the IR-predicted normal cells showed normal pap smear and seventy-one percent of the IR-predicted malignancies were confirmed by biopsies. Similar study was conducted on a larger population of 436 patients by Cohenford et al (1997) and achieved a sensitivity of 79% and specificity of 77% in the diagnosis of cervical cancer [158]. In the same year, a sensitivity of 98.6% and specificity of 98.8% in detecting cervical dysplasia was reported by Fung et al (1997) [159].

Subsequently, Chiriboga and co-workers (1997, 1998) conducted a series of studies to comprehensively understand the IR spectral features of cervical tissues and exfoliated cells [160-162]. In their first study, they collected IR spectra from cervical tissue slices using FTIR microscope [160]. CIN tissues showed unique glycogen bands at 1028 and 1151 cm^{-1} as compared to normal squamous tissues. Then, to better interpret the spectra of exfoliated cells, IR spectra were acquired from various layers of normal squamous tissues, including superficial layer, intermediate layer, parabasal layer, basal layer and stroma [161]. Lastly, they reported that IR spectra of exfoliated cells can be fitted by the spectra of mucus and various layers so that the contribution of different layers can be estimated [162].

Wood and et al (1998) took endeavor to reveal the potential confounding variables which may interfere or deteriorate the discrimination of neoplastic or malignant cells based on IR band intensity [163]. The results showed that saline, Leukocytes, severe infestations of *C. albicans*, fibroblast, endocervical mucins, sperm contamination and thrombocytes may affect IR measurement. Consequently, multivariate statistic technique was proposed for developing diagnostic algorithm. Cohenford et al (1998) found that normal-appearing cells from patients with dysplasia or malignancy are IR spectroscopically differentiable from normal cells of healthy patients [164]. Romeo et al (2005) suggested that benign changes and endocervical cells also influence the diagnosis of cervical cancer and therefore should be removed prior to PCA-based diagnosis [165]. They also mentioned that hormonal status led to a variation in spectral features of cervical cells; however, it did not confound the diagnosis of cervical dysplasia by PCA [166]. They also attempted to remove the blood component from cervical smear [167]. However, the discrimination of dysplasia from normal cells became even worse after removing leukocytes, which raised the question on the true

origins of spectral difference between normal and malignant cells.

Apart from the spectral analysis above, mapping and imaging were proposed to analyze cervical tissue cryo-section. Chang and colleagues (2003) created the images of tissue sections with the ratio of the integrated intensities under IR bands of 1130~1180 cm^{-1} and 1180~1260 cm^{-1}, allowing a discrimination of cervical dysplasia from normal tissue [168]. In the following year, cluster analysis was used to build the image of tissue sections so as to extract IR spectral features of different cell types and differentiate among different tissue pathologies. Bambery et al (2004) mapped tissue cryo-section with both IR band intensity and clusters generated by unsupervised HCA [169]. The results showed that the spectra from the same layer were almost resembled into the same cluster. In a separate report, Wood and co-workers (2004) distinguished successfully the layer structures of normal, LGSILs and HGSILs tissues with the same method and differentiated dysplasia from normal tissue [170]. IR bands at 1470~1740 cm^{-1} (amide I and II) showed a good correlation with tissue anatomy and histology. Steller et al (2006) suggested that IR spectroscopy coupled with Fuzzy C-means clustering for data reduction and HCA for classification have the ability to distinguish different tissue types as well as tissue pathologies [171].

To optimize the performance of IR spectra for screening cervical dysplasia and cancer, neural networks was introduced for developing diagnostic algorithms. Mark et al (2004) reported their success in grading cervical neoplasia (i.e., CIN 1, 2 and 3) with the use of probabilistic neural networks (PNN) with an accuracy rate of ~85% [172]. Podshyvalov et al (2005) differentiated cancer from normal tissues with an accuracy (>95%) [173].

1.6 Thesis Motivations, Objectives and Organization

1.6.1 Motivations and Objectives

NIR Raman spectroscopy exhibites advantages for the detection of cervical precancer and cancer, such as non-invasive, real-time and higher accuracy, as compared to those conventional screening/diagnosis methods (i.e., pap smear and colposcopy). On the other hand, it also holds advantages over other optical spectroscopic techniques (i.e., UV/VIS fluorescence and reflectance spectroscopy). For instance, it can yield higher specificity as it is a molecular fingerprint technique. Moreover, the use of NIR excitation light allows a longer penetration depth in tissue and consequently the acquisition of Raman signal originating in deeper layer of tissues [51]. In addition, NIR light is noncarcinogenetic. However, most of Raman studies in characterizing cervical tissues are limited in fingerprint region [98, 100, 102]. Note that the fiber-optics materials can produce Raman and fluorescence signal in fingerprint region which may interference the tissue Raman signal. Consequently, optical filtering modules are used to eliminate the interference from fiber-optics materials and as a result complicate the design of Raman probe. Recently, HW Raman spectroscopy has been proposed for tissue characterization [8]. HW Raman probe design can be simplified with the use of single fiber for both excitation and collection but no optical filtering module as fiber-optics materials do not produce Raman scattering in HW region [174]. Till now, HW Raman spectroscopy for cervical tissue discrimination has not been reported. Meanwhile, most of Raman studies for cancer diagnosis in various organs, including cervix, are conducted alone and separated from other optical techniques. Note that different optical techniques may provide insights into different aspects of tissue premalignant or malignant changes [175, 176]. In particular, for cervix, only limited studies were conducted to combine UV/VIS fluorescence and

reflectance for the diagnosis of cervical dysplasia [148]. Therefore, this thesis work is conducted towards two aims. The first one is to explore the potential of HW NIR Raman spectroscopy for the detection of cervical precancer *in vivo*. The second one is to investigate the feasibility of combining different optical spectroscopic techniques to further improve the diagnosis of cervical precancer as compared to the diagnosis by individual spectroscopic method alone. The specific aims are as following:

(1) To verify the capacity of NIR Raman spectroscopy in fingerprint region (800~1800 cm^{-1}) for discriminating cervical dysplasia *ex vivo*;

(2) To develop a rapid and portable NIR Raman spectroscopy system for *in vivo* Raman measurements on cervix. An *in vivo* fiber-optic Raman probe design with the use of ball lens is proposed to enhance the depth-selective Raman measurement on the cervix;

(3) To explore the feasibility of NIR Raman spectroscopy in HW region (2800~3700 cm^{-1}) for the *in vivo* diagnosis of cervical dysplasia with the use of the *in vivo* Raman system developed;

(4) To combine Raman and NIR autofluorescence background signal to improve the detection of cervical dysplasia as compared to the diagnosis by Raman spectroscopy alone;

(5) To optimize the optical diagnosis of cervical dysplasia through combining NIR Raman, UV/VIS fluorescence and diffuse reflectance spectroscopy.

1.6.2 Thesis Organization

This chapter has provided the background knowledge relevant to this thesis work, the objectives and motivations of this thesis work. The remaining thesis is organized as

following. Chapter 2 presents an *ex vivo* study on verifying the ability of Raman spectroscopy for differentiating benign, LGSILs and HGSILs of cervix. It describes a new Raman spectrum acquisition program integrated with an online-processing function developed for Raman spectral acquisition, spectral analysis and multivariate statistical analysis. Chapter 3 describes the development of a rapid portable NIR Raman spectroscopy for *in vivo* Raman measurement on cervix. It describes the design of a ball lens Raman probe for enhancing *in vivo* depth-selective Raman measurement on cervix in detail. MC simulation method was used to comprehensively evaluate the performance of this ball-lens Raman probe configured with a variety of ball lens' diameter and refractive index, and probe-tissue distance. Chapter 4 reviews our work on investigating the feasibility of HW Raman spectroscopy for the *in vivo* diagnosis of cervical precancer using the system depicted in Chapter 3. The results demonstrated the efficacy of the in-house developed Raman system and the potential of HW Raman for discriminating cervical precancer from normal tissues. Chapter 5 depicts our work on exploring the possibility of combining Raman and NIR autofluorescence to improve the diagnosis of cervical precancer. Chapter 6 reports our attempt to combine Raman spectroscopy with UV/VIS fluorescence and diffuse reflectance spectroscopic techniques for optimizing the optical diagnosis of cervical precancer. Chapter 7 summarizes the whole thesis and also gives the future directions.

Chapter 2 NIR Raman Spectroscopy for *Ex Vivo* Detection of Cervical Precancer: Multivariate Statistical Analysis and Spectral Modeling

Raman spectroscopy is a unique analytic probe for molecular vibrations and can provide molecular fingerprint information of biochemical compositions in tissues. It has proven to have the potential for diagnosing precancer and cancer in various human organs through probing the biochemical changes associated with tissue cancerous transformation. In this chapter, we presented our work on verifying the capabilities of Raman spectroscopy in fingerprint region (800~1800 cm^{-1}) for the *ex vivo* detection of cervical dysplasia. A rapid NIR Raman spectroscopy system combined with a bifurcate fiber-optic Raman probe was employed for the Raman spectroscopic measurements on cervical tissues. A novel data acquisition program was developed using LabVIEW and Matlab. In total, 80 spectra were acquired from 30 samples, including 24 spectra from 9 benign tissues, 34 spectra from 12 LGSILs tissues and 22 spectra from 9 HGSILs tissues. Empirical analysis method (i.e., intensity ratio) was used to differentiate precancer from benign tissues. PCA and LDA were employed to develop diagnostic algorithms. Moreover, tissue spectrum was fitted with reference spectra of pure biochemicals using linear least-squares fitting method to elucidate the biochemical changes behind Raman spectral variation associated with tissue dysplastic changes. The results demonstrated that NIR Raman spectroscopy has the potential to diagnose cervical precancer.

2.1 Materials and Methods

2.1.1 Cervical Tissue Samples

A total of 30 cervical biopsied tissue specimens were obtained from patients who underwent colposcopy-guided biopsy or surgical resection for clinically suspicious lesions or histologically proven premalignancies. All the patients preoperatively signed an informed consent permitting our Raman study on the tissues and this study was also approved by the Ethics Committee of the National Healthcare Group of Singapore. After biopsy or surgical resection, tissue specimens were immediately stored in physiologic saline solution (pH=7.4) and delivered to our laboratory for Raman spectroscopic measurement without any pretreatment on the tissue specimens. After Raman spectroscopic measurement, tissue specimens were fixed in 10% formalin solution and then submitted to the pathology laboratory for histopathological examination. The histology results showed that 9 specimens are benign, 12 specimens are LGSILs and 9 specimens are HGSILs as shown in table 2.1. Benign tissues comprise chronic cervicitis, squamous metaplasia, inflammation and koilocytic atypia.

Table 2.1 Statistics of tissue samples and Raman measurements.

Histology	Race				Age (mean)	Number of Biopsies	Number of Spectra
	Chinese	Indian	Malay	others			
Benign	9	0	0	0	29~56 (42.2)	9	24
LGSILs	7	0	2	1	27~64 (43.5)	12	34
HGSILs	6	1	1	1	22~48 (37.7)	9	22

2.1.2 Reference Spectra of Biochemicals

Five biochemicals were selected to fit Raman spectra of cervix. The five biochemicals are: collagen type I, glycogen, DNA, oleic acid and cholesterol (Sigma-Aldrich, St.

Louis, MO).

2.1.3 Raman Instrumentation

A rapid NIR Raman spectroscopy system was employed for the tissue Raman measurement in optical bioimaging laboratory at National University of Singapore. Figure 2.1 shows the schematic diagram of the Raman system [78]. The whole system mainly consists of a spectrum stabilized 785-nm diode laser (maximum output, 300 mW; B&W Tek, Newark, DE), a transmissive imaging spectrograph (Holospec f/2.2, Kaiser Optical Systems, Ann Arbor, MI), a NIR-optimized, back-illuminated, and deep-depletion CCD camera (1340×400 pixels at 20×20 μm/pixel; Spec-10: 400BR/LN, Princeton Instruments, Trenton, NJ), and a specially developed bifurcated fiber-optic Raman probe. The probe is 1.8 mm in outer diameter and 2.5 m in length, and comprises 33 ultralow-OH fused-silica fibers (NA=0.22) in which one 200-μm central fiber is used for laser light delivery while thirty-two 200-μm surrounding fibers are for Raman signal collection. The excitation light is exported out from the diode laser by a 200-μm core diameter fiber and then coupled into the central excitation fiber of the fiber-optic probe through an in-line filter module integrated with a narrow band-pass filter (LL01-785, Semrock, Rochester, NY) for suppressing laser noise, fluorescence, and Raman emission from the exporting fiber. Another narrow band-pass filter (centered at 785 nm, FWHM = ±2.5 mm) is coated on top of the distal end of the excitation fiber to reduce most of the fused-silica noise generated in the excitation fiber before the excitation light hits the tissue. The back-scattered Raman photons are filtered by an edge long-pass filter (cut off at 800 nm) coated on top of the distal end of the collection fibers to block the elastically scattered excitation light prior to being collected by the surrounding fibers. Subsequently, the Raman photons are coupled into the fiber bundle adapter via an in-line filter module integrated with an edge long-pass

filter (LP02-785RU, Semrock, Rochester, NY) for further reduction of the scattered laser light while permitting the scattered-tissue Raman photons to pass through toward the Raman spectrograph. Eventually, the fiber bundle feeds the Raman photons into the Raman spectrograph. The fiber bundle comprises sixty-four 100-μm ultralow-OH fibers (NA=0.22; length=1.5 m) and is packed in a round geometry at the long-pass in-line filter module. At the entrance slit of the Raman spectrograph, the fiber bundle is arranged as a parabolic linear array in an orientation opposite to the image aberration of the transmissive spectrograph, which effectively corrects spectrograph image aberrations, thereby enables the hardware binning of the entire CCD chip to improve the SNR up to 20-fold (400/√400) compared to the complete software binning. The spectral resolution of our system is approximately 9 cm^{-1} (a 100-μm core diameter fiber covers 6 pixels with a 1.25×magnification of the spectrograph; each pixel covers 1.5 cm^{-1}).

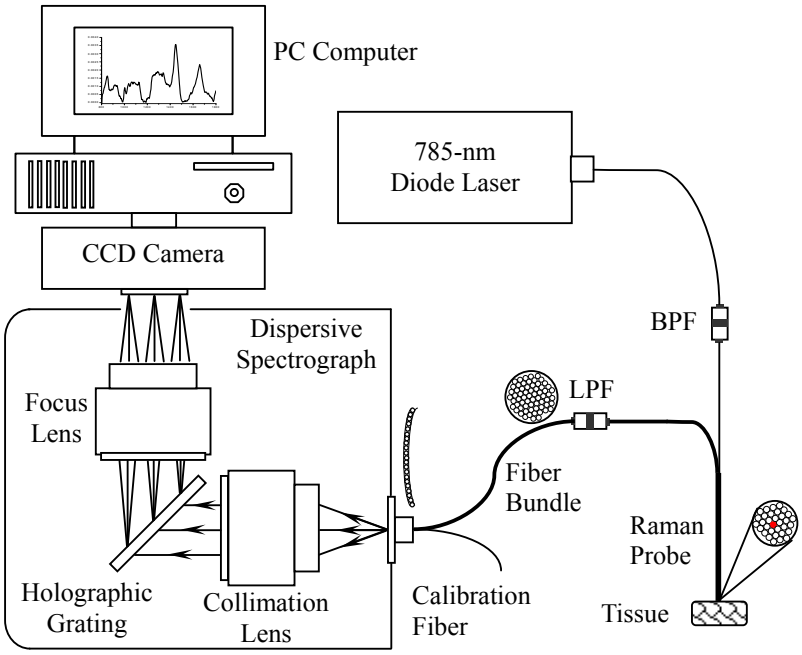


Figure 2.1 Schematic of the NIR Raman spectroscopy system. BPF: Band Pass Filter; LPF: Long Pass Filter.

2.1.4 Raman Data Acquisition Program Development

The commercial software for hardware controlling and data acquisition is WinSpec (Princeton Instruments, Trenton, NJ). In addition to data acquisition, it is also integrated with correction and math functions (i.e., background subtraction and smoothing). However, it cannot extract the true tissue Raman spectrum from the raw spectrum by eliminating the tissue autofluorescence background. Therefore, we developed our own Raman data acquisition program using LabVIEW (National Instruments, Austin, TX) and Matlab (The MathWorks, Natick, MA). The most significant function provided by our program is to fit the autofluorescence background with a polynomial and then subtract it from raw tissue spectrum to yield the true tissue Raman spectrum alone in real-time. The program interface is shown in Fig. 2.2. The left two spectral windows are a mixture of raw spectrum and fitted autofluorescence background (top) and Raman spectrum after autofluorescence background subtraction (bottom). The right two spectral windows show the counterpart after smoothing. The

complete functions integrated into this program include: (1) set detector temperature; (2) set integration time; (3) wavelength-dependent system response function calibration; (4) dark subtraction; (5) smoothing; (6) autofluorescence background fitting and subtraction; (7) data save.

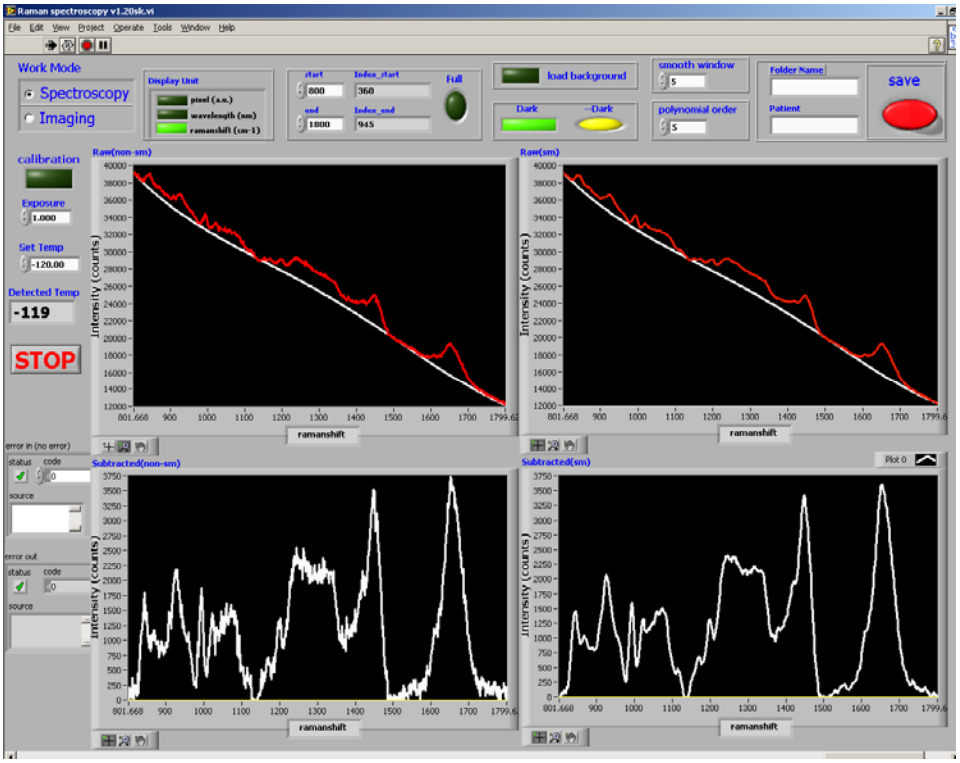


Figure 2.2 The interface of Raman data acquisition program developed using LabVIEW and Matlab.

2.1.5 Raman Measurement

All the tissue Raman measurements were done in a light-tight box which can eliminate most of the interference from the environmental lights (e.g., computer monitor, light indicators). Tissue specimens were mounted onto aluminum foil which does not produce Raman signal interference. The laser light irradiance on tissue is 1.56 W/cm^2 which is less than the American National Standards Institute (ANSI) maximum permissible skin exposure limit for a 785-nm laser light beam [177]. All the Raman spectra were acquired over the range of 781.1~934.1 nm. The spectrum acquisition

time was adjusted from 1 to 3 seconds according to the overall intensity and SNR. Raman frequencies were calibrated with the spectra of cyclohexane, acetone and barium sulfate to an accuracy of ±2cm^{-1} (the digital data resolution of the CCD detector). All spectra were corrected for the wavelength-dependent intensity response function of the system using a standard lamp (RS-10, EG&G Gamma Scientific, San Diego, CA).

2.1.6 Data Preprocessing

The raw spectra acquired from tissue were truncated in the range of 800~1800 cm^{-1} for preprocessing and analysis. The Raw spectra were composed of intense autofluorescence, noise and very weak tissue Raman scattering signal. Firstly, the raw spectra were preprocessed by a first order Savitzky-Golay filter (window size: 5 pixels) to reduce noise [178]. Secondly, a fifth order polynomial was found to optimally fit the autofluorescence background signal of all the noise-reduced spectra [51]. The fitted polynomial was subtracted from the noise-reduced spectrum to yield the Raman spectrum alone. Finally, each background-subtracted tissue Raman spectrum was normalized to the integration area under the spectrum to correct the inter- and/or intra-patient variation in absolute spectral intensity so as to enable a better comparison on the spectral line-shape and relative Raman band intensity. Prior to the statistical analysis, empirical analysis method (i.e., intensity ratio of different Raman bands) was employed to explore the diagnostic potential of Raman spectroscopy for cervical precancer. The histology results serve as the gold standard for us to evaluate the sensitivity and specificity of the diagnostic algorithm.

2.1.7 Multivariate Statistical Analysis

PCA was firstly performed on the tissue spectral dataset to reduce the data dimension

while retaining the most diagnostically significant information for effective tissue classification. To eliminate the influence of inter- and/or intra-subject spectral variability on PCA, the entire spectra were standardized so that the mean of the spectra was zero, and the standard deviation (SD) of all the spectral intensities was one. Mean centering ensures that the principal components (PCs) form an orthogonal basis [179, 180]. The standardized dataset was assembled into data matrices with wavenumber columns and individual case rows. Thus, PCA was conducted on the standardized spectral data matrices to generate PCs comprising a reduced number of orthogonal variables that accounted for most of the total variance in original spectra. Each loading vector is related to the original spectrum by a variable called the PC score, which represents the weight of that particular component against the basis spectrum. PC scores reflected the differences between different classes. One-way analysis of variance (ANOVA) was used to identify the most diagnostically significant PCs (p<0.1) among benign, LGSILs and HGSILs. These significant PC scores were lastly selected as the input of LDA for developing trichotomous classification algorithms. LDA determines the discriminant function that maximizes the variances in the data among different groups while minimizing the variances among the members within the same group. The performance of the diagnostic algorithms rendered by the PCA-LDA models for correctly predicting the tissue groups was estimated in an unbiased manner using the leave-one spectrum-out, cross-validation method on all model spectra [179, 181].

To evaluate the overall trichotomous classification accuracy, a Mossman's three-way Receiver Operating Characteristic (ROC) testing was performed [182]. A ROC surface was created by utilizing the posterior probabilities belonging to each of the three tissue groups (i.e., benign, LGSILs and HGSILs) produced by PCA-LDA diagnostic model.

Each case has three probabilities P1, P2 and P3 (P1+P2+P3=1) belonging to group 1, 2 and 3, respectively. A threshold value (e.g., T-value 1 = 0:0.1:1) is chosen for judging whether a case is from group 1 or not. If P1>T-value 1, then this case is clustered into group1; If P1<T-value 1, another threshold (e.g., T-value 2 = −1: 0.2:1) is chosen for deciding whether this case is from group 2 or group3. If (P2−P3)>T-value 2, then this case is classified into group 2; otherwise, it is classified into group 3. As a result, with 11×11 different combinations of T-value 1 and 2, 121×3 matrices of correct classification rates will be yielded, of which each column is the correct classification rate for one group. Then, this matrix can be plotted in three-dimension to create the ROC surface. The volume under the three-way ROC surface (VUS) equals the probability that the tissue is identified correctly. Thus, VUS represents the overall three-group classification accuracy as the area under ROC curve for a two-group classification.

2.1.8 Spectral Modeling

Linear least-squares fitting with nonnegativity constraints was employed to fit tissue spectrum with reference spectra of biochemicals (i.e., collagen, glycogen, DNA, cholesterol and oleic acid) [17]. The contribution of each reference spectrum to tissue spectrum was calculated by normalizing the fit coefficients such that they sum to one. The normalized coefficients represent the percentage of each component's contribution to the target tissue spectrum. LDA was used to evaluate the diagnostic value of the normalized fitting coefficients for all the five reference spectra.

2.2 Results

2.2.1 Spectral Feature Analysis

Figure 2.3 shows the averaged Raman spectra yielded by subtracting the

autofluorescence background from the raw spectra. It was apparent that the tissue Raman spectrum was dominated by a few prominent Raman peaks located in the vicinities of 849, 932, 1004, 1039, 1063, 1085, 1125, 1178, 1206, 1254, 1285, 1317, 1339, 1449, 1580 and 1658 cm^{-1}. However, the tissue Raman spectrum was subject to a significant inter- and/or intra- patient intensity variation illustrated by SD as shown in the Fig. 2.3. Moreover, it is also challenging in clinical practice to standardize the absolute intensities of spectral measurements. Therefore, in this study, we precluded tissue diagnosis based on absolute intensities alone; instead, we focused on the analysis of the spectral line shape and relative Raman band intensity by normalizing the Raman spectrum to the integrated area under the spectrum.

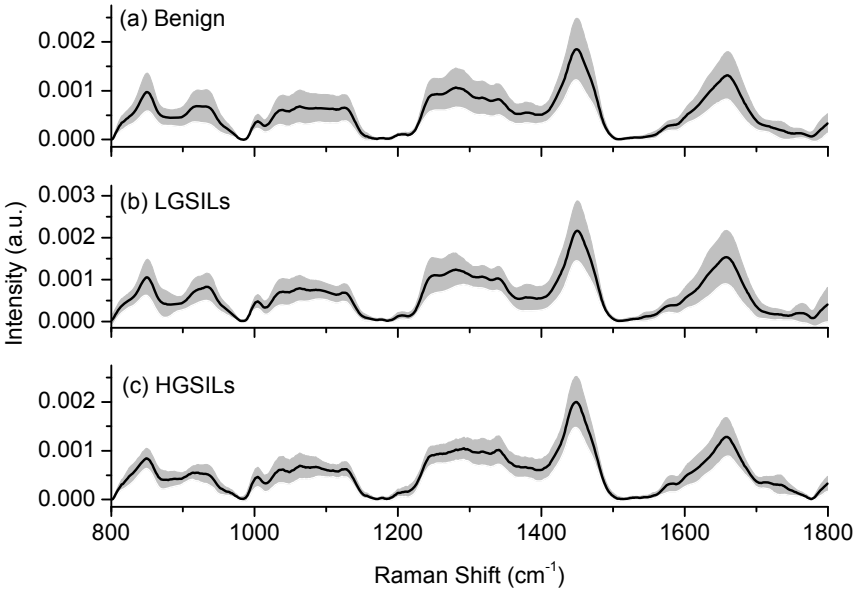


Figure 2.3 The averaged Raman spectra±1SD of: (a) benign=24, (b) LGSILs=34, and (c) HGSILs=22. Line: averaged spectrum; Grey band: ±1SD.

Figure 2.4 depicts the average intensity-normalized Raman spectra: (a) benign and LGSILs, (c) benign and HGSILs, (e) LGSILs and HGSILs. Significant spectral differences were observed among different types of tissues (i.e., benign, LGSILs and HGSILs). To better visualize the spectral difference, the difference spectra were produced, including LGSILs−benign, HGSILs−benign and HGSILs−LGSILs as shown

in Figs. 2.4((b), (d) and (f)), respectively. The spectral intensity was found to increase or decrease progressively from benign to LGSILs to HGSILs in the spectral region filled with grey color. For instance, Raman peaks at 1004, 1339 and 1449 cm^{-1} showed a continuous increase in intensity with the progression of tissue changes from benign to HGSILs while Raman peaks at 849 and 1658 cm^{-1} showed otherwise. It was noted that those prominent Raman peaks can be tentatively assigned to chemical bonds in tissues according to the literature report [5, 51]. Hence, those Raman spectral changes can represent the biochemical composition and molecular structural changes and therefore be related to the tissue precancerous changes. Eventually, they may provide useful information at molecular level for tissue characterization.

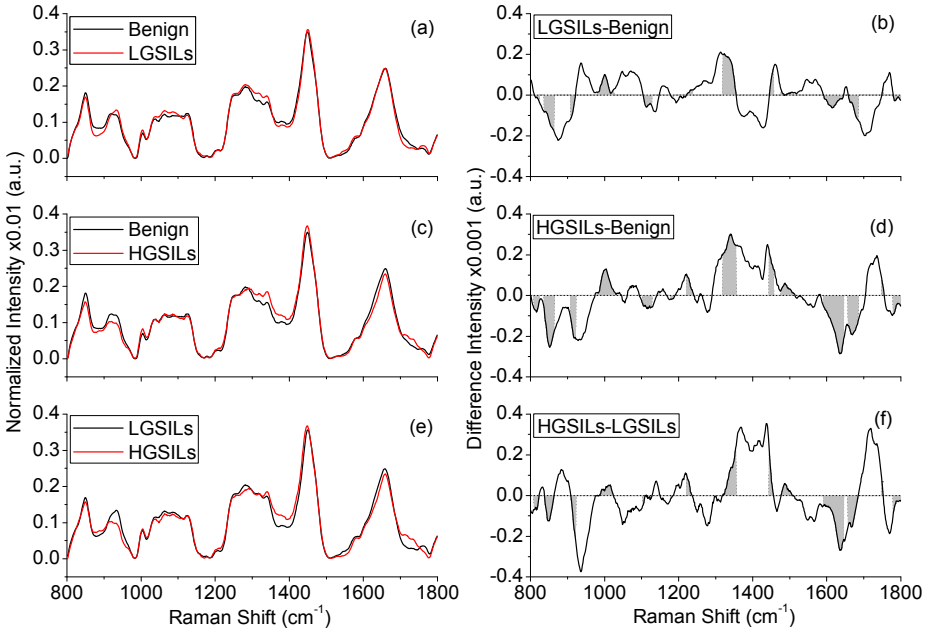


Figure 2.4 The averaged normalized Raman spectra of: (a) benign and LGSILs, (c) benign and HGSILs and (e) LGSILs and HGSILs. The corresponding difference spectra are: (b) LGSILs−benign, (d) HGSILs−benign, and (f) HGSILs−LGSILs.

2.2.2 Empirical Analysis

To verify the potential of Raman spectroscopy for tissue classification, empirical method (i.e., intensity ratio of Raman peaks) was firstly employed to develop diagnostic functions for characterizing the dysplastic cervical tissues by utilizing those

prominent Raman peaks. Figure 2.5 shows several scatter plotting examples of the intensity ratios of Raman peaks for the differentiation among benign, LGSILs and HGSILs with high accuracies. For example, single ratio of I_{849}/I_{1004} can achieve a sensitivity of 73.5% (25/34) and specificity of 79.2% (19/24) in differentiating LGSILs from benign tissues with the diagnostic function of ratio (2.56). For the discrimination of HGSILs from benign tissues, the combinations of I_{932}/I_{1449} *vs* I_{1339}/I_{1658} and I_{932}/I_{1449} *vs* I_{1449}/I_{1658} were found to be optimal, showing a sensitivity of 100.0% and specificity of 66.7% (16/24) and 70.8% (17/24), respectively. For the separation of HGSILs from LGSILs, the combinations of I_{932}/I_{1254} *vs* I_{932}/I_{1658} and I_{932}/I_{1449} *vs* I_{1004}/I_{1658} achieved the best diagnostic performance, showing a sensitivity of 68.2% (15/22) and 63.6% (14/22), and specificity of 97.1% (33/34) and 100.0% (34/34), respectively.

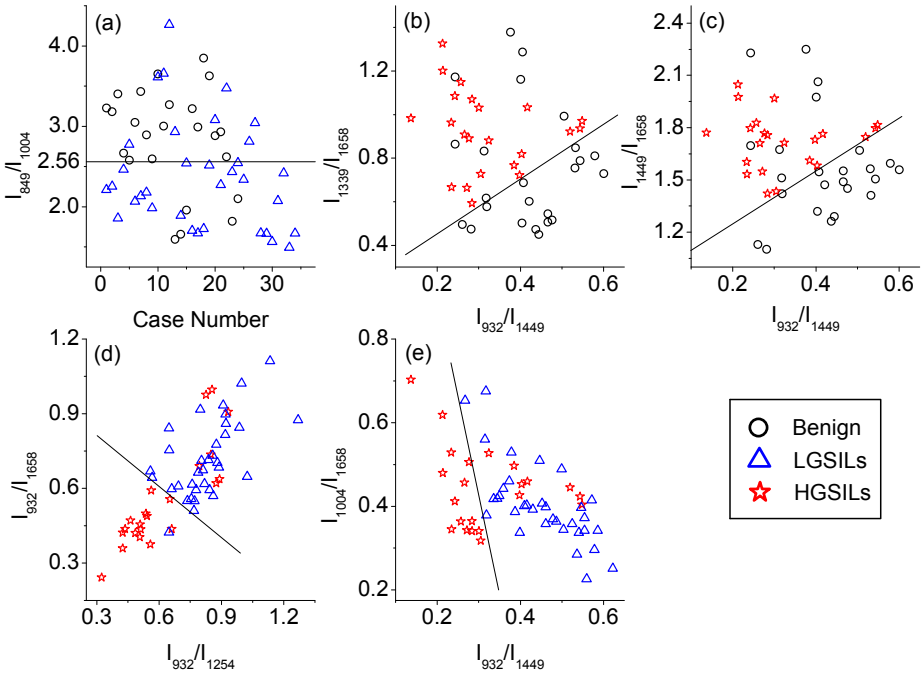


Figure 2.5 Scatter plots of the intensity ratio of Raman bands: (a) benign *vs* LGSILs, I_{849}/I_{1004}; (b) benign *vs* HGSILs, I_{932}/I_{1449} *vs* I_{1339}/I_{1658}; (c) benign *vs* HGSILs, I_{932}/I_{1449} *vs* I_{1449}/I_{1658}; (d) LGSILs *vs* HGSILs, I_{932}/I_{1254} *vs* I_{932}/I_{1658}; (e) LGSILs *vs* HGSILs, I_{932}/I_{1449} *vs* I_{1004}/I_{1658}; Simple straight-line diagnostic function can achieve sensitivities and specificities of: (a) 73.5% (25/34) and 79.2% (19/24); (b) 100.0% (22/22) and 66.7% (16/24); (c) 100.0% (22/22) and 70.8% (17/24); (d) 68.2% (15/22) and 97.1% (33/34); (e) 63.6% (14/22) and 100.0% (34/34), respectively. Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

2.2.3 PCA-LDA and ROC Analysis

It was noticed that the diagnostic accuracy achieved by intensity ratio method is limited. This may be explained by the use of only limited spectral variables, indicating a limited use of the Raman spectral features. Therefore, a comprehensive use of all the Raman spectral variables was desired to further improve the diagnostic performance. Hence, multivariate statistical techniques (i.e., PCA and LDA) were proposed to develop the diagnostic algorithms with the utilization of the spectral features in the whole spectral range (800~1800 cm^{-1}).

Figure 2.6 displays the first six diagnostically significant PC loadings yielded by PCA on the tissue Raman spectral data: (a) PC1, (b) PC2, (c) PC4, (d) PC5, (e) PC6 and (f) PC11. It was found that all the six PCs appeared to contain many narrow bands and some of the PC features can roughly be related to the Raman spectra with peaks and troughs positions similar to those of tissue Raman spectra. Table 2.2 lists those six PCs' mean scores±1SD for benign, LGSILs and HGSILs tissue groups, *p*-value and percent of total variance. The total variance covered by PC1, PC2, PC4, PC5, PC6 and PC11 is 27.2%, 15.3%, 11.5%, 8.0%, 4.4% and 1.3%, respectively. The first three PCs (PC1, PC2 and PC5) can represent more than 50% of the total variance. These results indicate the amount of the diagnostic information contained in the Raman spectra for differentiating among benign, LGSILs and HGSILs tissues.

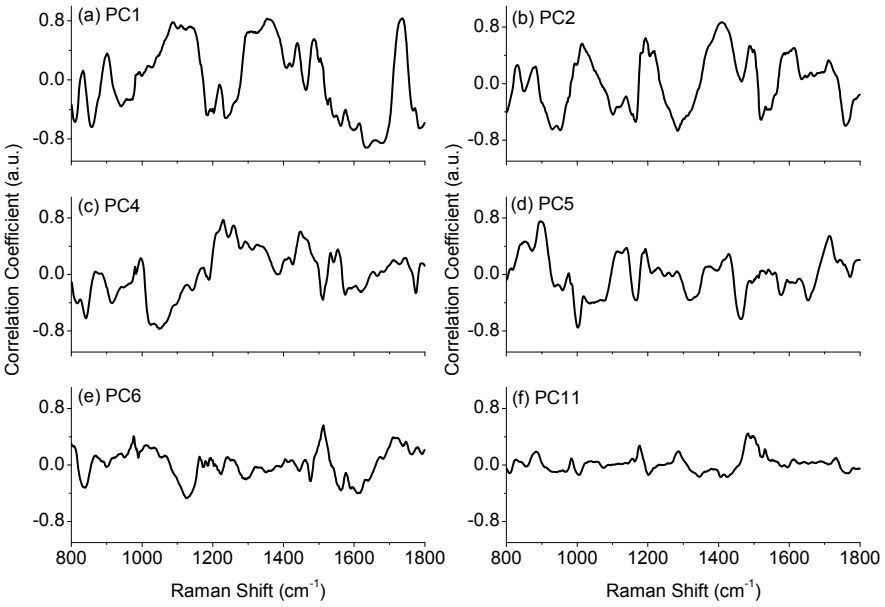


Figure 2.6 Examples of the first six diagnostically significant PCs with p-value<0.1: (a) PC1, (b) PC2, (c) PC4, (d) PC5, (e) PC6 and (f) PC11.

Table 2.2 The mean scores±1SD for benign, LGSILs and HGSILs tissue groups, p-value, percent of total variance of the first six diagnostically significant PCs.

PC	Mean Score ±SD			p-value	Percent of Total Variance
	Benign	LGSILs	HGSILs		
PC1	-0.187±0.908	-0.146±1.150	0.429±0.718	0.05876	27.2
PC2	0.077±0.759	-0.411±0.936	0.552±1.075	0.00125	15.3
PC4	-0.223±1.045	-0.119±1.066	0.428±0.712	0.05603	11.5
PC5	0.566±0.779	-0.530±0.942	0.203±0.916	0.00004	8.0
PC6	-0.122±0.998	-0.167±0.991	0.391±0.952	0.09574	4.4
PC11	0.095±0.972	0.253±0.889	-0.495±1.056	0.01859	1.3

To illustrate the ability of PC scores for tissue characterization, the first six diagnostically significant PC scores were scatter-plotted as shown in Figs. 2.7 to 2.9. Figure 2.7 describes the scatter plots of PC2, PC4, PC5 and PC11 for the separation between benign and LGSILs: (a) PC2 *vs* PC5, (b) PC2 *vs* PC11 (c) PC4 *vs* PC5, (d) PC5 *vs* PC11. It was noticed that LGSILs can be discriminated from benign tissues by straight-line diagnostic functions. Almost 75% of the tissues can be characterized correctly, indicating that PC2, PC4, PC5 and PC11 contained a wealth of diagnostic

information for discriminating LGSILs from benign. Similarly, Fig. 2.8 depicts the scatter plots of PC1, PC2, PC4, PC5 and PC11 for distinguishing HGSILs from benign tissues: (a) PC1 *vs* PC11, (b) PC2 *vs* PC4 (c) PC2 *vs* PC11, (d) PC4 *vs* PC5. The straight-line can roughly represent the diagnostic function which can distinguish around 76% of the tissues, indicating the diagnostic information provided by the PCs. Figure 2.9 shows the scatter plots of PC2, PC4, PC5 and PC11, which can separate HGSILs from LGSILs with an accuracy of about 80%, demonstrating the diagnostic ability of those PCs.

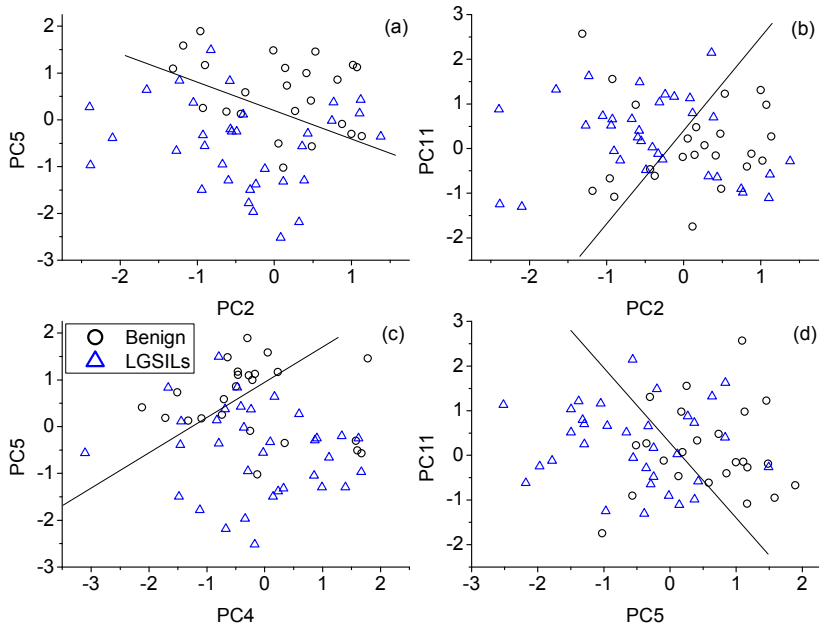


Figure 2.7 Scatter plots of the diagnostically significant PC scores for benign and LGSILs tissues: (a) PC2 *vs* PC5, (b) PC2 *vs* PC11 (c) PC4 *vs* PC5, (d) PC5 *vs* PC11. LGSILs can be discriminated from benign tissues by straight-line diagnostic functions: (a) PC5 = −0.61×PC2+0.2, (b) PC11 = 2.1×PC2+0.41, (c) PC5 = 0.76×PC4+0.95 and (d) PC11 = −1.68×PC5+0.29. The corresponding sensitivities and specificities are: (a) 79.4% (27/34) and 75.0% (18/24), (b) 76.5% (26/34) and 70.8% (17/24), (c) 85.3% (29/34) and 70.8% (17/24), and (d) 79.4 % (27/34) and 75.0 % (18/24), respectively. Key: (○ in black) benign; (Δ in blue) LGSILs.

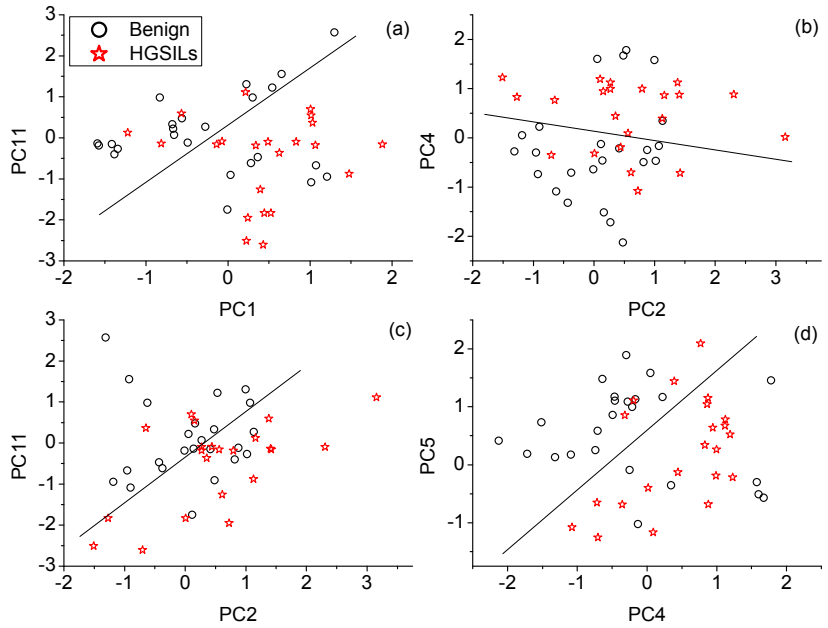


Figure 2.8 Scatter plots of the diagnostically significant PC scores for benign and HGSILs tissues: (a) PC1 *vs* PC11, (b) PC2 *vs* PC4 (c) PC2 *vs* PC11, (d) PC4 *vs* PC5. HGSILs can be discriminated from benign tissues by straight-line diagnostic functions: (a) PC11 = 1.39×PC1+0.31, (b) PC4 = −0.19×PC2+0.14, (c) PC11 = 1.11×PC2−0.34 and (d) PC5 = 1.03×PC4+0.6. The corresponding sensitivities and specificities are: (a) 81.8% (18/22) and 70.8% (17/24), (b) 72.7% (16/22) and 79.2% (19/24), (c) 86.4% (19/22) and 70.8% (17/24), and (d) 81.8% (18/22) and 66.7% (16/24), respectively. Key: (○ in black) benign; (☆in red) HGSILs.

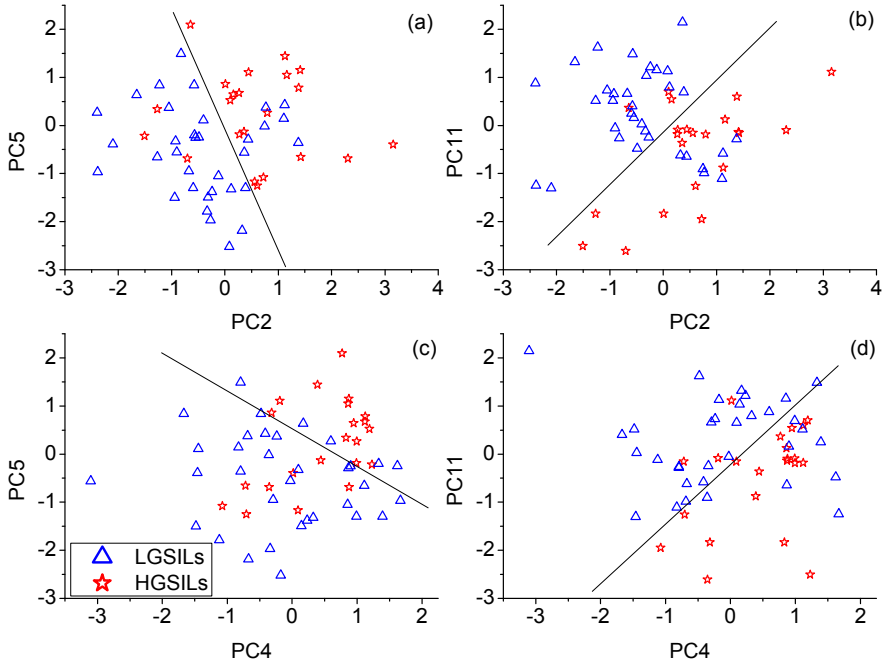


Figure 2.9 Scatter plots of the diagnostically significant PC scores for LGSILs and HGSILs tissues: (a) PC2 *vs* PC5, (b) PC2 *vs* PC11 (c) PC4 *vs* PC5, (d) PC4 *vs* PC11. HGSILs can be discriminated from LGSILs by straight-line diagnostic functions: (a) PC5 = −2.5×PC2−0.08,

(b) PC11 = 1.09×PC2−0.14, (c) PC5 = −0.78×PC4+0.53 and (d) PC11 = 1.24×PC4−0.21. The corresponding sensitivities and specificities are: (a) 86.4% (19/22) and 79.4% (27/34), (b) 86.4% (19/22) and 79.4% (27/34), (c) 63.6% (14/22) and % 85.3(29/34), and (d) 86.4% (19/22) and 79.4% (27/34), respectively. Key: (Δ in blue) LGSILs; (☆in red) HGSILs.

To further improve the diagnosis, LDA was used to incorporate all the significant PCs for developing diagnostic algorithms. For the three-group classification, two linear discrimination (LD) functions were generated. Figure 2.10 shows the scatter plot of two LD weights for the classification of benign, LGSILs and HGSILs. It was seen that most of the tissues can be clustered into the three histopathological groups (i.e., benign, LGSILs and HGSILs). The posterior probabilities of each tissue site belonging to benign, LGSILs and HGSILs were plotted in a ternary plate (Fig. 2.11). It was found that all HGSILs had posterior probabilities higher than 80% to be HGSILs except two HGSILs with posterior probabilities between 50% and 80%. In comparison, posterior probabilities of LGSILs and benign distributed more diversely. Seven benigns exhibited posterior probabilities below 75% (i.e., 0~25% (2), 25~50% (2) and 50~75% (3)) belonging to benign. Five LGSILs' posterior probability was located in 0~25% (1), 25~50% (2) and 50~75% (2) belonging to LGSILs. This distribution also indicates the efficacy of the PCA-LDA model for characterizing benign, LGSILs and HGSILs. Table 2.3 summarizes the classification results achieved by the PCA-LDA diagnostic algorithms with leave-one-out cross-validation methods. Overall, 92.5% of the tissues were predicted correctly, among which 83.1% of benign, 94.1% of LGSILs and 100.0% of HGSILs (sensitivity) were discriminated correctly. Four benigns were misclassified, of which two as LGSILs and two as HGSILs. Meanwhile, two LGSILs were misclassified as benign and HGSILs, respectively. The corresponding specificity and accuracy were 98.2% and 93.8% for benign, 96.6% and 95.0% for LGSILs, 94.8% and 96.3% for HGSILs. Figure 2.12 depicts the three-way ROC surface for this three-group classification with the ROC-VUS of 0.815.

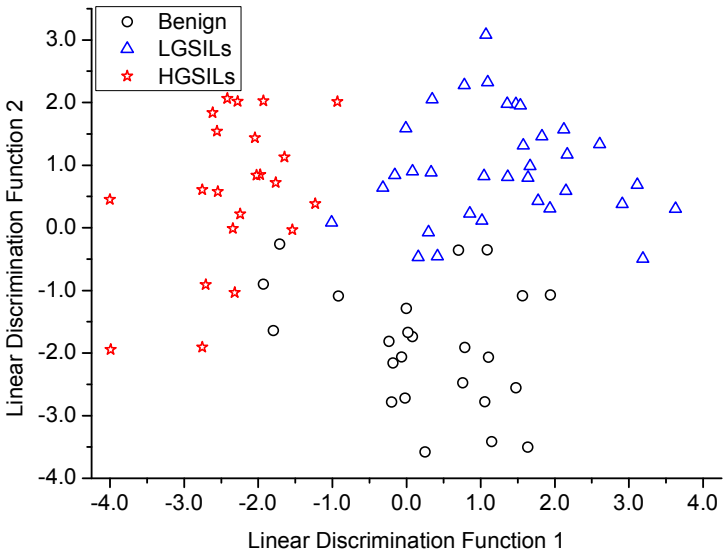


Figure 2.10 Scatter plot of two LD function weights for benign, LGSILs and HGSILs tissues tested with leave-one-out cross-validation. Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

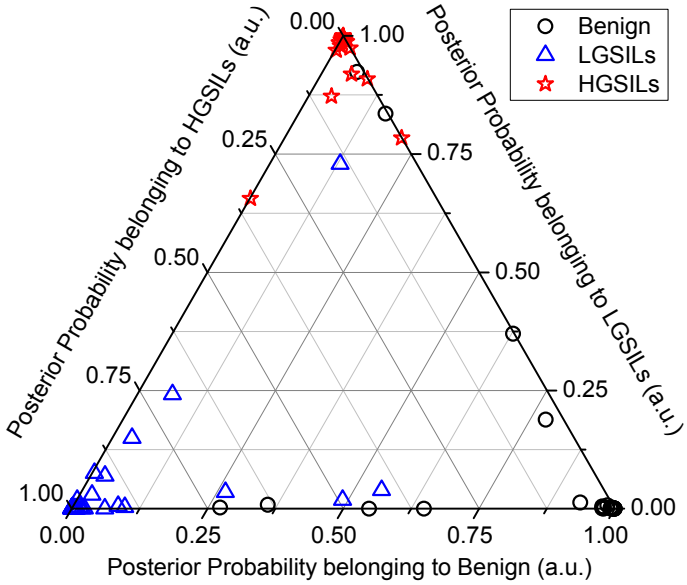


Figure 2.11 Two-dimensional ternary plot of the posterior probabilities belonging to benign, LGSILs and HGSILs, respectively, using the PCA-LDA-based spectral classification with leave-one spectrum-out, cross-validation method. Each vertex of the triangle represents a 100% confidence that the tissue is benign, LGSILs or HGSILs. Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

Table 2.3 Classification results of Raman-prediction for the three cervical tissue groups yielded by the PCA-LDA diagnostic algorithms tested with leave-one-out cross-validation method.

	Raman-predicted Tissue Type			
		Dysplasia		
Tissue Type	Benign	LGSILs	HGSILs	Total
Benign	20	2	2	24
LGSILs	1	32	1	34
HGSILs	0	0	22	22
Outcome measure*	benign	LGSILs	HGSILs	
Sensitivity (%)	83.3	94.1	100.0	
Specificity (%)	98.2	96.6	94.8	
Accuracy (%)	93.8	95.0	96.3	92.5

*Outcome measure for each group is calculated against the remaining two tissue groups.

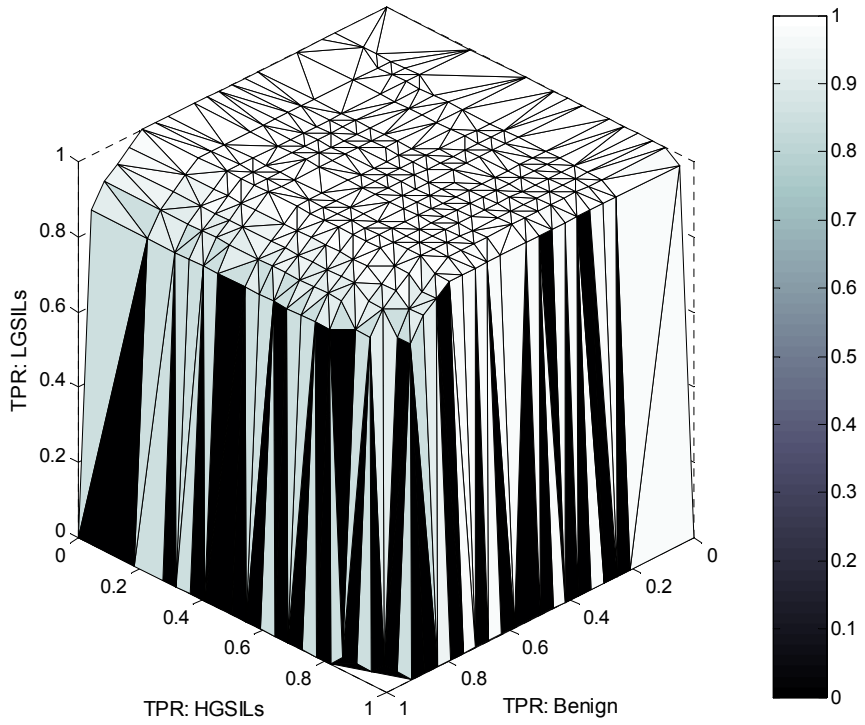


Figure 2.12 Three-dimensional view of the ROC surface calculated from the posterior probabilities belonging to benign, LGSILs and HGSILs with a VUS of 0.815.

2.2.4 Biochemical Model of Tissue Spectrum

To explore the biochemical origins accounting for the Raman spectral features, tissue spectrum was fitted with reference spectra of pure biochemicals using linear least-squares method with nonnegativity constraints. Figure 2.13 shows the Raman spectra of reference biochemicals, including cholesterol, DNA, oleic acid, collagen,

and glycogen which are considered as the major Raman scatterers in cervical tissues [173, 183-186]. It was obvious that each biochemical showed unique Raman spectral pattern composed of several Raman peaks. To estimate the similarity between each reference spectrum and tissue spectrum, correlation was calculated for all the biochemicals with tissue spectra. The correlation coefficients are yielded by dividing the covariance of biochemical spectrum and tissue spectrum by the standard deviations of biochemical spectrum and tissue spectrum. The yielded correlation coefficients were cholesterol-0.62, oleic acid-0.52, DNA-0.82, collagen-0.88 and glycogen-0.41. The correlation coefficients quantified the similarity between reference spectrum and tissue spectrum, indicating that collagen was the most alike with tissue spectrum.

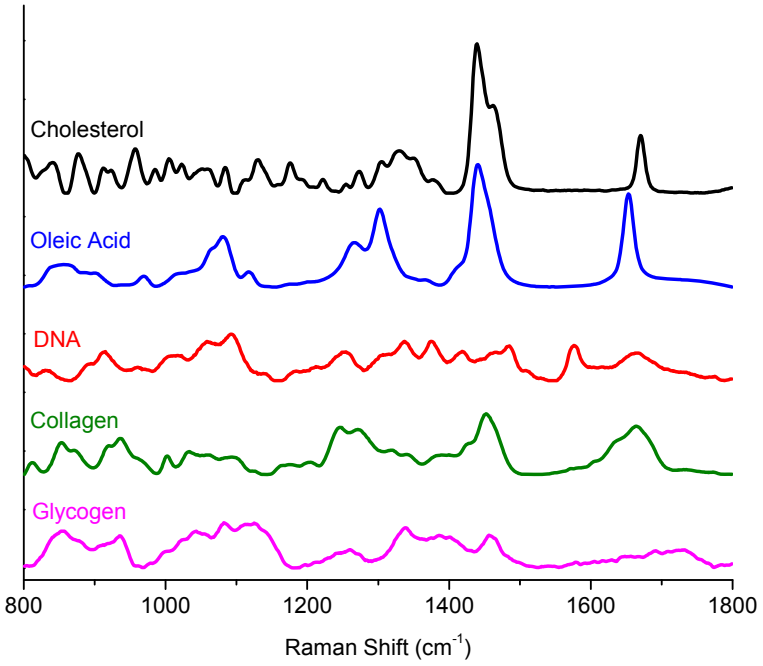


Figure 2.13 Reference Raman spectra of glycogen, collagen, DNA, oleic acid and cholesterol.

Figure 2.14 shows an example of tissue spectrum, fitted spectrum and the corresponding residue: (a) benign tissue, (b) dysplasia tissue. It was noticed that tissue spectrum was fitted well with the five reference spectra except the spectral regions of 1000~1200 cm^{-1}. Figure 2.15 shows the mean fitting contribution±1SD of each

reference biochemical to benign and dysplasia tissues. It was found that collagen made the largest contribution to Raman signal of both benign and dysplasia tissues. Moreover, the contribution of each reference biochemical was significantly different between benign and dysplasia tissues illustrated by unpaired two-sided Student's *t*-test. For example, collagen and glycogen exhibited a significantly reduced contribution to Raman signal of dysplasia tissue as compared to benign tissue. In contrast, cholesterol, DNA and oleic acid showed a significantly higher contribution to Raman signal of dysplasia than benign tissue. The above results demonstrate that Raman spectroscopy is able to directly assess the biochemical changes associated with premalignant transformation in the cervix. To further evaluate the diagnostic value of the fitting coefficients, LDA was implemented on the fitting coefficients to develop diagnostic algorithm. It yielded a sensitivity, specificity and total accuracy of 69.6%, 75.0% and 71.3%, respectively.

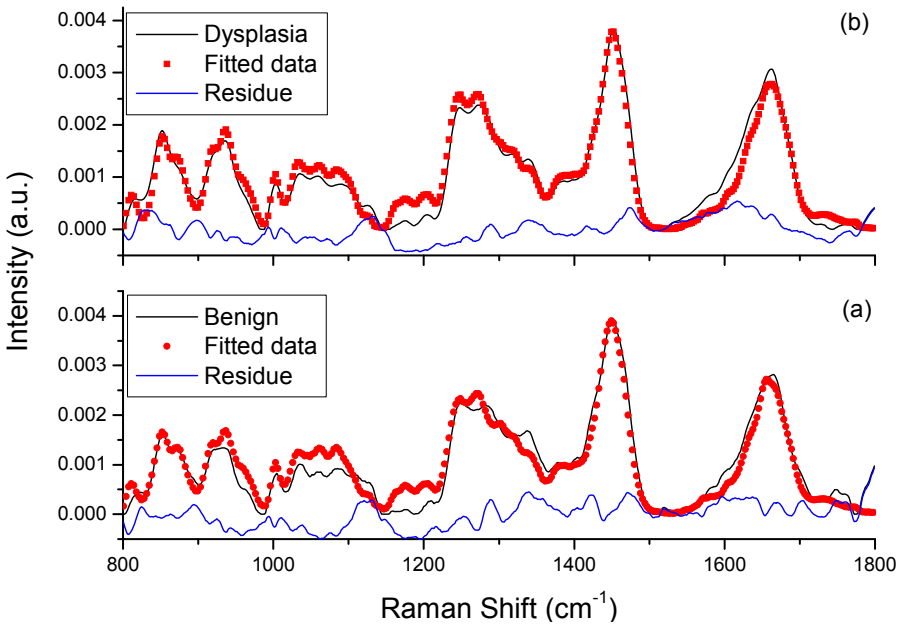


Figure 2.14 Comparison of tissue spectrum and fitted spectrum with reference spectra: (a) benign and (b) dysplasia. Residue is produced by subtracting fitted spectrum from tissue spectrum.

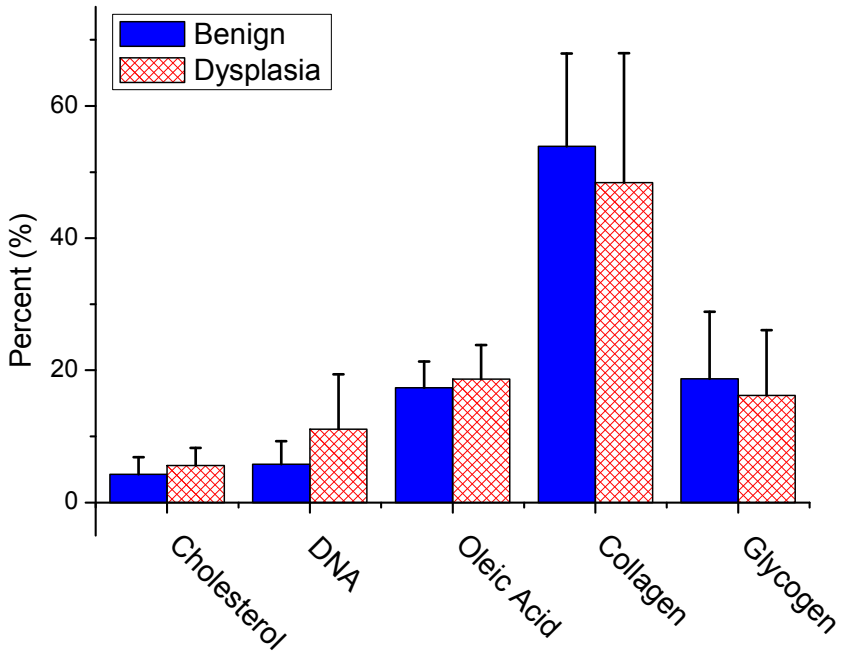


Figure 2.15 Mean normalized fitting coefficients with 1±SD for component biochemicals, including glycogen, collagen, oleic acid, DNA and cholesterol for benign and dysplasia tissues.

2.3 Discussion

In this study, we verified the capacity of NIR Raman spectroscopy for characterizing pre-dysplasia (benign changes: chronic cervicitis, squamous metaplasia, inflammation and koilocytic type atypia) and dysplasia (LGSILs and HGSILs) in the cervix. In the clinic diagnosis point of view, the discrimination between benign and dysplasia is not the whole colposcopic practice in clinic. The discrimination between LGSILs from HGSILs is also clinically desirable. The reason is that LGSILs have a high rate to regress to normal without any treatment and however the current colposcopic diagnosis is likely to overestimate LGSILs as HGSILs [187]. Meanwhile, discrimination between benign and LGSILs is even more challenging because some benign changes of cervix may also show some features of LGSILs [188]. For instance, inflammation may also exhibit vascular atypia which is the hallmark of high grade lesion, and metaplasia may exhibit acetowhite due to the cellular changes [187, 188].

This may increase the false-negative rate for diagnosing dysplasia. Therefore, a three-class diagnostic model based on PCA-LDA was developed to discriminate either of benign, LGSILs and HGSILs from the remainder.

Good quality Raman spectra were acquired successfully from benign, LGSILs and HGSILs tissues. A high spectral similarity occurred among Raman spectra of benign, LGSILs and HGSILs, of which all showed almost the same major Raman vibration bands in the vicinities of 849, 932, 1004, 1039, 1063, 1085, 1125, 1178, 1206, 1254, 1285, 1317, 1339, 1449, 1580 and 1658 cm^{-1}. This implies that the composition of the major biochemicals did not alter a lot with the presence of different pathologies. Consequently, only subtle spectral differences existed among benign, LGSILs and HGSILs, which were manifested clearly in the difference spectra as shown in Fig. 2.4 ((b, (d) and (f)). To elucidate the biochemical basis behind the spectral differences above, those major Raman bands were tentatively assigned to vibration bonds of a variety of biochemicals in tissues as shown in table 2.4 according to literature [5, 51, 108, 189]. Consequently, the spectral variations among different types of tissues can be explained by the corresponding biochemical changes in tissue accompanying the tissue neoplastic changes. For example, Raman peak at 1658 cm^{-1} (Amide I, C=O stretching mode of collagen and elastin) appeared to decrease as tissue progresses from benign to LGSILs to HGSILs, which is consistent with the reported findings in literature [98, 100, 190]. This may be due to a decrease in the relative concentration of collagen and elastin. This may also be attributed in part to the epithelium thickening and increased cellular density due to the enlarged nucleus-to-cytoplasm ratio associated with tissue malignant transformation. This induces a larger attenuation of both excitation light and Raman signal from the collagen in stroma since the underlying stroma is well recognized to be rich in collagen and elastin [70, 72, 191, 192]. In addition, the Raman

signal originating from the stroma of dysplastic cervix also suffers from larger hemoglobin absorption due to the neovascularization as compared to normal or benign cervix [193]. Meanwhile, another two peaks at 849 (C-C stretching of proline ring and ring breathing mode of tyrosine) and 1125 cm^{-1} (C-C stretching of lipid, C-N stretching of protein) followed the same variation trend. The intensity decrease in Raman peak at 849 cm^{-1} suggests that the relative concentration of proline and tyrosine drop in neoplastic cervical tissues. For the Raman peak at 1125 cm^{-1}, it may result from the epithelium thickening. In contrast, some other Raman peaks at 1004, 1339 and 1449 cm^{-1} varied in the opposition way. Raman peak at 1004 cm^{-1} which is associated with phenylanaline showed a higher signal in precancer tissue, indicating an increase in the essential amino acid phenylanaline relative to the total Raman-active components in the precancer tissue due to the fact that cancer cells take up more essential amino acids [194]. Raman peak at 1339 cm^{-1} is mainly attributed to nucleic acid. Hence, its intensity elevation may be attributed to the increase in the concentration of nucleic acid relative to the total Raman-active scatterers, which gets in an agreement with the literature [5] as well as the well-accepted histopathological feature of increased nucleus-to-cytoplasm ratio in malignant tissue [195]. For the Raman peak at 1449 cm^{-1}, the intensity increase is probably because the lipid content is elevated in neoplastic tissue due to the increases in metabolic activities [185, 186].

Table 2.4 Tentative assignments of major Raman vibration bands present in the Raman spectra of cervical tissues [5, 51, 108, 189].

Peak positions (cm^{-1})	Biochemical assignments
849	C−C stretching of proline ring, Ring breathing mode of tyrosine
932	C−C stretching mode of proline and valine and protein backbone (α−helix conformation)/glycogen
1004	Symmetric ring breathing mode of phenylalanine
1039	C−H in-plane bending mode of phenylalanine
1063	C−N stretching mode of protein
1085	C−C/C−O stretching of phospholipids
1125	C−C stretching of lipid, C−N stretching of protein
1178	C−H in-plane bending mode of tyrosine
1206	$C-C_6H_5$ stretching mode of tryptophan and phenylalanine
1254	Amide III (C−N stretching mode of proteins, indicating mainly α−helix conformation)
1285	Amide III
1317	CH_3CH_2 twisting mode of collagen/lipids
1339	CH_3CH_2 wagging mode of collagen/nucleic acid
1449	CH_2 and CH_3 bending mode of collagen, CH_2 bending, scissoring mode of phospholipids
1580	C=C bending mode of phenylalanine, pyrimidine ring
1658	Amide I (C=O stretching mode of collagen and elastin, α−helix conformation)

The diagnostic potential of these prominent Raman bands were explored directly by intensity ratio method. For the differentiation between benign and LGSILs, ratio of Raman bands I_{849}/I_{1004} achieved an accuracy of 75%, implying that the changes of proline ring and tyrosine (I_{849}) and phenylanaline (I_{1004}) is diagnostically valuable. As for separating HGSILs from benign, higher-than-80% accuracy achieved by ratios of Raman bands at 932, 1340, 1449 and 1658 cm^{-1} implies the diagnostic value of the changes of proline, DNA, lipid and protein. Among them, proline, lipid and protein combined with Amide III can also be used to discriminate HGSILs from LGSILs with an accuracy of higher than 85%. These results demonstrate that Raman spectral variation among different types of tissues can represent the biochemical composition

and molecular structure changes associated with tissue malignant transformation, and contain a wealth of information for tissue diagnosis.

However, the diagnostic accuracy yielded by the intensity ratio method is limited due to the fact that only a few individual Raman bands are included for developing diagnostic algorithm while missing a lot of biochemical information contained in the spectra unused [5]. Therefore, multivariate statistical techniques (e.g., PCA) are desired to utilize all the subtle spectral changes to differentiate different types of tissues. PCA has been widely used to retrieve the Raman spectral features into the size-reduced dataset. Moreover, the PC scores have proven to have the potential for distinguishing tissues under different histopathological conditions by simple scatter plots of the PC scores. Teh et al achieved a good classification of gastric dysplasia from normal gastric tissue by using the scatter plot of a pair of PCs with a accuracy of about 90% [70]. This method has also been verified by Huang et al [175, 190]. Similarly, in this study, the scatter plots of the first six significant PCs prove again to be able to cluster the tissues into different histopathological categories (i.e., benign, LGSILs and HGSILs).

However, scatter plot of PC scores only used two PC scores while missing the remaining diagnostic PC scores. Hence, in this study, LDA was further employed to utilize all the diagnostic features covered by all the significant PCs. ANOVA was used to find out those diagnostically significant PC scores as the input of LDA for developing diagnostic functions. The PCA-LDA diagnostic results show that all the HGSILs tissues can be differentiated from both benign and LGSILs. Only two of LGSILs were misclassified, of which one was classified as benign and the other as HGSILs. Four of benign tissue sites were misdiagnosed as LGSILs (2) and HGSILs (2), indicating a relatively low specificity.

Furthermore, to better understanding the biochemical alterations with tissue neoplastic transformation, we also modeled the tissue spectra with reference spectra of pure biochemicals, including lipids (i.e., cholesterol and oleic acid), nucleic acid (i.e., DNA), protein (i.e., collagen) and carbohydrates (i.e., glycogen). This tissue spectral modeling method has been used for uncovering the biochemical composition alteration with tissue malignancy behind Raman spectral variations [16, 60, 61, 196]. The results show that collagen made the largest contribution to tissue spectrum. This is attributed to the fact that collagen takes up a larger proportion of cervix as compared to the rest. Meanwhile, it was observed that both collagen and glycogen exhibited a reduced contribution to dysplastic tissue spectrum than benign tissue spectrum. The reduced contribution from collagen is confirmed by the decreased intensity of Raman peak at 1658 cm^{-1} in precancer, which has been explained above. The reduced level of glycogen is probably due to the increased cell proliferation [173, 183]. In contrast, DNA, cholesterol and oleic acid showed an elevated contribution to dysplastic tissue spectrum. DNA is representative of the nucleic acids and one of the major DNA Raman peaks located at 1339 cm^{-1} also exhibited an increased intensity. The increased level of DNA is suggestive of an enlarged ratio of nuclei to cytoplasm [184]. The elevated level of oleic acid in precancer tissue may contribute to the continuous growth stimulus [185]. As for cholesterol, it is an essential element of cell membrane and plays a key role in the modulation of membrane fluidity. The increased level of cholesterol in dysplastic cervix may be related to the higher minimum surface tension of dysplastic cervical tissue [186]. From the discussions above, spectral modeling of tissue spectrum with reference spectra of pure biochemicals is an effective tool for identifying the origins of biochemical changes with tissue dysplastic transformation.

2.4 Conclusion

In this study, we succeeded in measuring good quality NIR Raman spectra from excised human cervical tissues within a few seconds. Significant spectral differences were observed among benign, LGSILs and HGSILs tissues. Both empirical analysis and tissue characterization with PCA-LDA model demonstrated that NIR Raman spectroscopy has the ability to identify benign, LGSILs and HGSILs. Although the current dataset size is small, the results still make it promising to apply NIR Raman spectroscopy to *in vivo* clinic diagnosis so as to overcome the limitations of the conventional screen methods (i.e., pap smear and colposcopy). In addition, spectral modeling of tissue spectrum with reference spectra of pure biochemicals can be used to elucidate the origins of biochemical changes behind Raman spectral changes.

Chapter 3 *In Vivo* NIR Raman Spectroscopy Development for the Detection of Cervical Precancer

NIR Raman spectroscopy has been extensively applied to probe the histopathological changes in a variety of human organs. In the early times, NIR Raman spectral measurement was mostly limited to be *ex vivo* due to the technical challenges. In recent years, the technical advances, especially in diode laser, CCD detector and fiber optics, paved the way for *in vivo* NIR Raman spectroscopic measurement on human tissues. However, there still exist some weaknesses, including large size, long integration time and interference from the materials of fiber optics. This chapter reviews my work on developing a rapid and portable NIR Raman spectroscopy system coupled with a ball lens fiber-optic Raman probe for the *in vivo* Raman measurement on epithelial tissues (i.e., cervix). Monte Carlo (MC) simulation was used to evaluate the Raman probe design with various configurations (i.e., diameter and refractive index of ball lens). Experimental evaluation of a ball lens Raman probe design on a two-layer tissue phantom confirms the potential of ball lens Raman probe design for efficient depth-selected measurement on epithelial tissue.

3.1 Introduction

At the early time of Raman spectroscopy for biological application, NIR FT Raman spectroscopy was accepted as the main form of Raman spectroscopy. NIR FT Raman spectroscopy combined with fiber-optic assemblies has proven sensitive enough for Raman spectroscopic measurement on biological samples [197]. Subsequently, it was employed to study Raman scattering properties of human tissues *in situ*, such as lens [198, 199], artery [200], breast [201] and colon [27]. In principle, FT Raman spectroscopy is a multiplex technique, which modulates a series of wavelengths with an interferometer. The modulation generates a complex interferogram which is subsequently monitored by a single-channel detector. FT Raman spectroscopy shows several advantages compared to dispersive Raman spectroscopy. Firstly, it can avoid the fluorescence interference from the biological sample since it uses NIR light as excitation source while conventional dispersive Raman spectroscopy uses UV/VIS light. This is very beneficial especially for biological samples which emit strong fluorescence under the excitation of UV/VIS light, masking the weak tissue Raman signal. To date, 1064-nm laser light has become the most common excitation light for FT Raman spectroscopy. Secondly, it can achieve high spectral resolution stemming from the interferometric effect. In contrast, the spectral resolution of dispersive Raman spectroscopy is limited by the slit width. If reducing the slit width, the throughput will be reduced although the spectral resolution is increased. Lastly, it can acquire the spectrum in broad wavelength region than most of the dispersive Raman spectroscopy without sacrificing the high spectral resolution.

However, FT Raman spectroscopy also subjects to some drawbacks, which lead it to be used in limited areas. The first one is long integration time. Note that Raman cross-section is inversely proportional to the excitation light wavelength. It means that

Raman signal will become weaker with the excitation light at longer wavelength. Moreover, FT Raman spectroscopy scans Raman signal at one by one wavenumber, which consequently lengthens the acquisition time. The long integration time impedes the application of FT Raman in those areas, which require rapid measurement. Secondly, it suffers from water absorption, especially when measuring samples in aqueous media. The overtone region of water absorbs not only the excitation light but also the Raman scattering due to C-H stretch modes of samples. Besides, the water absorption makes reproducible temperature measurement difficult [202]. In addition, FT Raman spectroscopy has large body size. All those weaknesses above hinder its application on human tissue *in vivo*.

With the advance in CCD, diode laser and fiber optics, NIR dispersive Raman spectroscopy has undergone a rapid progress in the past two decades, exhibiting a great potential for the detection of cancer and precancer in various human organs *ex vivo* and *in vivo*. In this study, for *in vivo* Raman measurement on cervix in clinic, we aimed to develop a rapid dispersive Raman spectroscopy system with good SNR and compact body size. Figure 3.1 shows the schematic of the dispersive Raman spectroscopy. The whole system consists of: (1) laser source for Raman scattering excitation; (2) spectrometer for collecting Raman signal to create Raman spectrum; (3) Raman probe for excitation light delivery and Raman signal collection; and (4) computer for controlling spectrometer, Raman spectrum display and save. Hence, the development of this Raman spectroscopy system mainly includes the following procedures as below:

(1) Determine the wavelength and select proper laser source;

(2) Spectrometer selection;

(3) Raman probe design;

In the following sections, the procedures above will be described and discussed in detail.

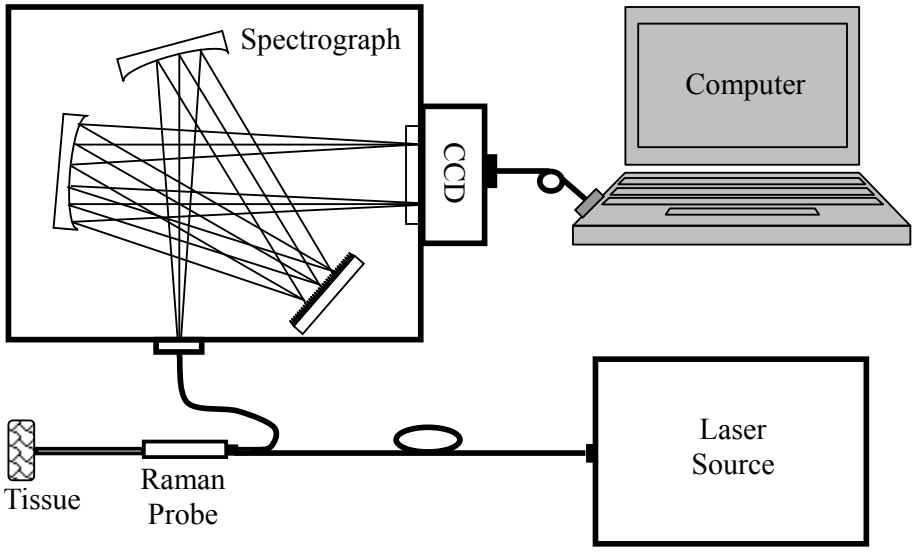


Figure 3.1 Schematic of the developed dispersive Raman spectroscopy.

3.2 Excitation Light Source

To date, almost all the Raman spectroscopy employs laser light as excitation light. Thereby, laser light selection is one of the key parts of Raman spectroscopy development. The factors which should be considered in laser selection are wavelength, output power and cost. At the same time, laser light selection must be done together with the design of other parts of Raman spectroscopy. For example, laser light wavelength determines the wavelength range of Raman scattering and hence it determines the grating central wavelength. Meanwhile, it also requires the detector to have high quantum efficiency in the wavelength range of Raman scattering. In the current market, there are a big variety of lasers which have been used as excitation light source for Raman spectroscopy as below [4]:

(1) Ar^+ and Kr^+ lasers are the two main forms of ion lasers. Ion laser can provide high power, and a variety of wavelengths from UV to NIR (Ar^+: 244, 257, 457, 488 and 514.5 nm; Kr^+: 406, 647 and 752 nm), relatively long lifetime, and hence was widely

used as excitation light for Raman spectroscopy in the early period of Raman spectroscopy. The output wavelength is very stable and accurately known since it is dependent on the particular atomic transition. However, the routine use of ion laser with Raman spectroscopy is limited by the high cost, complexity and cooling requirement.

(2) Nd:YAG is found to take the place of ion laser for its relatively high efficiency, compact size and no need of water-cooling. The Nd:YAG lasers at 532 and 1064 nm have been in very common use with Raman spectroscopy. Especially, 1064-nm Nd:YAG laser is almost the exclusive light source for FT Raman spectroscopy. However, the high cost prohibits its widespread use in Raman spectroscopy.

(3) He-Ne laser receives great attractiveness as excitation light source of Raman spectroscopy for its advantages of accurate wavelength and narrow output spectral width. However, the output power is low (~10 mW) and as a result long integration time is required for Raman measurement on human tissue showing weak Raman scattering. On the other hand, high output power causes He-Ne laser to be too large to be used in routine. Meanwhile, the most common laser light by He-Ne is at 632.8 nm, under which the fluorescence emission of human tissue is relatively high so that the SNR is reduced by the strong fluorescence-determined short noise.

(4) Ti:Sapphire lasers are advantaged by its robustness, tunability and high output power, and hence attractive for NIR Raman spectroscopy, especially for resonance Raman. However, it is subject to high cost.

(5) Diode lasers have been widely used instead of various other lasers above, especially for NIR Raman measurement on human tissues. Inasmuch, it shows advantages of good stability, small spectral band width, high power, compact size and

low cost.

Selection of excitation light wavelength is mainly dependent on several factors as below:

(1) Intensity of Raman scattering

(2) Fluorescence background

(3) Safety to human body

(4) Compatibility with other components of Raman system

Firstly, Raman scattering of biological tissues is inherently very weak. Therefore, great effort should be taken to increase the intensity of Raman scattering. In theory, Raman scattering intensity is proportional to Raman cross-section, which is inversely proportional to the forth power of the excitation light wavelength [4]. As a result, to increase the Raman scattering intensity, short wavelength is desired. However, shortening the excitation light wavelength (e.g., UV/VIS light) will lead to an increase in the tissue autofluorescence emission. The reason is that most of endogenous fluorophores are highly fluorescent under UV/VIS light excitation. Moreover, fluorescence intensity is usually 6~8 orders higher than Raman intensity. As a result, the relatively intense fluorescence signal will obscure the weak Raman signal. To reduce fluorescence interference, longer wavelength is preferred. Figure 3.2 shows the Raman spectra of human tissues excited by the common lasers used for Raman spectroscopy [4]. It was found that Raman spectra under 784- and 830-nm light excitation exhibit obvious Raman peaks. This suggests the use of NIR excitation light. Besides, considering the safety to human body, the use of short wavelength (e.g., UV light) is limited in clinic due to the mutagenic effect. Whereas, NIR light is favored as it causes no photochemical effect [175, 203]. Note that the Raman intensity decreases as the excitation light wavelength becomes longer and consequently a long spectrum

acquisition time is necessitated to obtain Raman spectrum with high SNR. However, the advance in NIR diode laser with high power and the multichannel-detector (e.g., CCD detector) with high sensitivity in NIR region allows the use of NIR light as the excitation source for Raman spectroscopy on biological tissues.

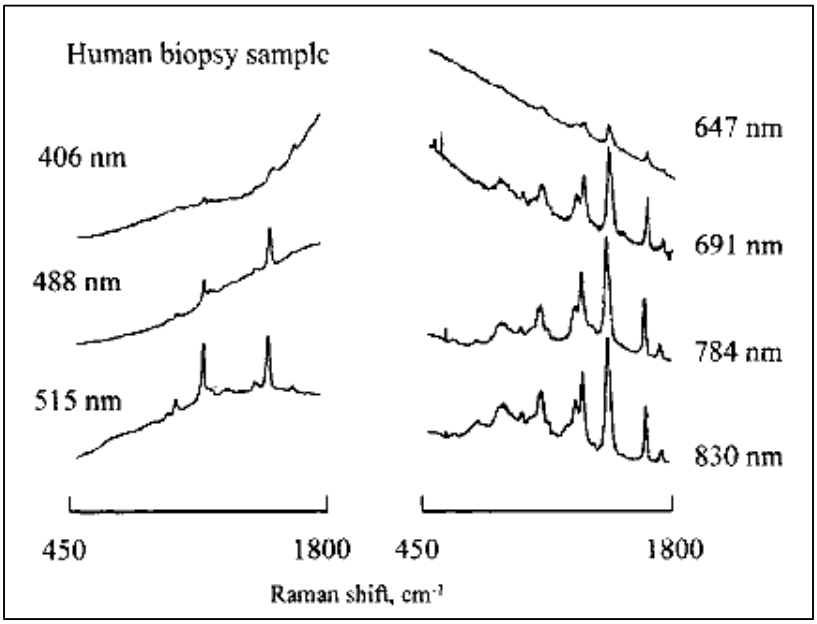


Figure 3.2 Raman spectra of human biopsied samples under different excitation lights (reprinted from Ref. [4]).

Till now, NIR laser light, especially 785- and 830-nm diode lasers, has been widely used as the excitation light source for dispersive Raman spectroscopy. Stone and coworkers (2002, 2003) employed 830-nm diode laser as excitation light to explore the potential of NIR Raman spectroscopy for the detection of epithelial cancer and precancer *ex vivo* [189, 204]. Anita and coworkers (1998, 2001, 2007, 2009) succeeded in using both 785- and 789-nm diode lasers for acquiring Raman spectrum from cervix *ex vivo* and *in vivo* [98, 100-102]. Our group (2001, 2003~2006, 2008~2010) has also demonstrated that 785-nm diode laser light enables an acquisition of good-quality Raman spectrum without much intense fluorescence interference from various epithelial tissues *ex vivo* and *in vivo*, including stomach, colon, lung, skin and cervix

[46, 51, 52, 67, 70-73, 78, 175, 205-207]. Therefore, 785-nm diode laser (B&W TEK, Newark, DE) is chosen as the excitation source for its small full width at half maximum (FWHM), good stability and high output power. The spectral width of the excitation line has a large impact on the spectral resolution of tissue Raman spectrum as Raman scattering is not dependent on wavelength but Raman shift which is the frequency shift relative to excitation light. For the same reason, the wavelength stability of excitation light is also very important as it causes a shift of Raman vibration bands due to the same biochemical bonds and consequently results in spectral variation with the lapse of time. The specification of this diode laser is shown in Table 3.1.

Table 3.1 Specification of BTK 785-nm diode laser.

Peak Wavelength	785±0.3 nm
FWHM	Typical<0.2 nm; Max<0.3 nm
Power Output	>100 mW
Optical Power Stability	<3%, p-p over 8 hours
Noise	<0.5% rms
Output Fiber	Multi-model fiber (100 μm, NA=0.22)
Connector	SMA 905
Expected Lifetime	>10,000 hours
Input Power	240V-AC
Control	TTL modulation 0-100 kHz
Cooling	Internal thermoelectric cooled
Warm-up Time	<15 minutes
Physical Dimension	265×155×100 mm
Weight	2 kg
Ambient Temperature	10-35 ºC
Humidity	5-95%, non-condensing

3.3 Spectrometer

Spectrometer is actually the core component of Raman spectroscopy system, which is an integration of spectrograph and detector. In particular, dispersive Raman spectroscopy mainly comprises a dispersive spectrograph and multi-channel detector. At present, dispersive spectrograph can be configured based on two different schemas, including Czerny-Turner structure and holographic transmission grating-based configurations as shown in Fig. 3.3. Czerny-Turner spectrograph (Fig. 3.3(a)) is the conventional form of dispersive spectrograph. In principle, the light is fed into the spectrograph via an entrance slit and subsequently is collimated by a curved collimating mirror. The collimated light is shed onto the diffraction grating and then reflected off at wavelength-dependent angle. Subsequently, the dispersed light is focused onto the detector. The light at different wavelengths is imaged on different columns of the array detector. Finally, the spectrum is created through binning the signal along each column. The spectral resolution is determined by the entrance slit width, groove density of the grating, column number and pixel size of the array detector. The other form of dispersive spectrograph is based on holographic transmission grating (Fig 3.3(b)). The light is focused onto the entrance slit and then collimated by lens. The collimated light is then dispersed by a transmission holographic grating. Subsequently, the dispersed light is focused onto the detector, and light at various wavelengths is imaged on the different columns of the detector. The spectrum is yielded through vertically binning the collected signal along each column. The major difference between Czerny-Turner and holographic transmission grating is the working mode which is off-axis and on-axis, respectively. On-axis working mode with optical axis normal to the lenses before and after grating allows an excellent optical aberration correction as well as good image quality. To date, low f number (i.e.,

f/1.8) can be achieved with this holographic grating design, allowing large throughput [208]. In contrast, dispersive spectrograph in off-axis mode is subject to astigmatism and other aberration, including spherical aberration and coma. To eliminate the aberration, small aperture is usually used in Czerny-Turner, which consequently limits the throughput of the spectrograph. In this study, Czerny-Turner spectrograph defeats holographic transmission spectrograph by virtue of its compact size, portability and low cost. The relatively low throughput can be compensated by the high sensitivity of detector, especially the advent of high performance CCD.

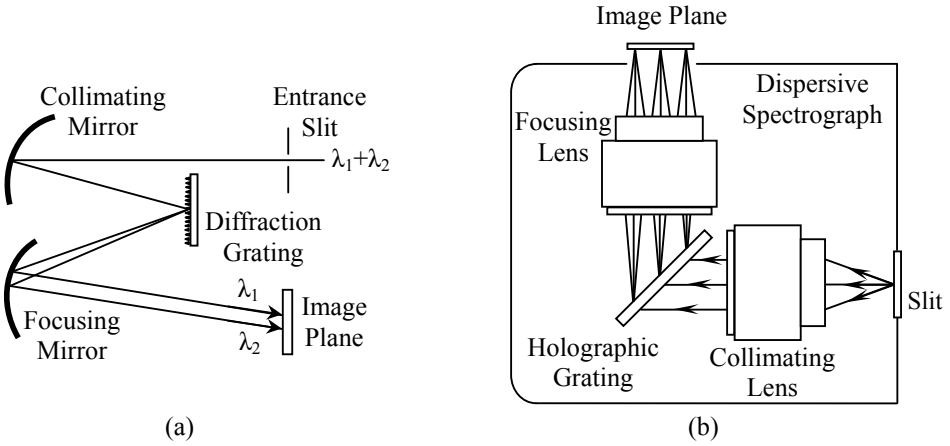


Figure 3.3 System schematic of dispersive spectrographs: (a) Czerny-Turner configuration; (b) Holographic transmission grating-based configuration.

In 1980s, CCDs emerged as a promising multichannel detector to facilitate the development of Raman spectroscopy, especially dispersive Raman spectroscopy. The first CCD was invented by George Smith and Willard Boyle in 1969 [209]. The fundamental concept is to store charge in potential wells created on the surface of a semiconductor and move the charge (representing information) over the surface by moving the potential minimum. CCD is sort of electronic memory that can be charged by light. It can hold a charge corresponding to variable shades of light, which makes them useful as imaging devices for cameras, scanners, and fax machines. Currently, the most common types of CCD as multi-channel detector in dispersive Raman

spectroscopy are Front-illuminated CCD (FI-CCD) and Back-illuminated CCD (BI-CCD, also called Back-thinned CCD) as shown in Fig. 3.4. FI-CCD means that photons are shed on the side of CCD containing the circuit mask, so called front side, while the other side is called back side. However, its quantum efficiency is limited (i.e., 40% for visible light) due to the big absorption of light by the gate oxide film, polysilicon electrodes, and surface protective film deposited on CCD front side. To improve the quantum efficiency, especially for the detection of low light, BI-CCD was developed, in which photons are shed on the back side of CCD and the generated electrons go through the CCD and are immobilized and stored in the potential well on the front side. Compared to FI-CCD, no film is needed on top of the back side of CCD so that the light loss in those films can be avoided. Consequently, BI-CCD has much higher quantum efficiency. As shown in Fig. 3.5, BI-CCD's quantum efficiency is almost 2.5 times higher in NIR region than FT-CCD's. However, BI-CCD suffers from etaloning effect due to the multi-reflections of the light between the front and back surfaces. This leads to constructive and destructive interference fringes which artificially modulate spectrum. This etaloning effect can be reduced or minimized by various ways: (1) NIR-optimized anti-reflection coating to reduce the light reflected back into CCD; (2) Increase the thickness of silicon; (3) The use of low f/# spectrometer; (4) Full-frame binning of CCD to create spectrum.

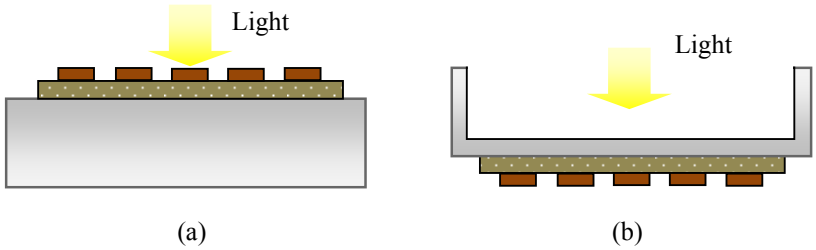


Figure 3.4 Schema of: (a) FI-CCD and (b) BI-CCD

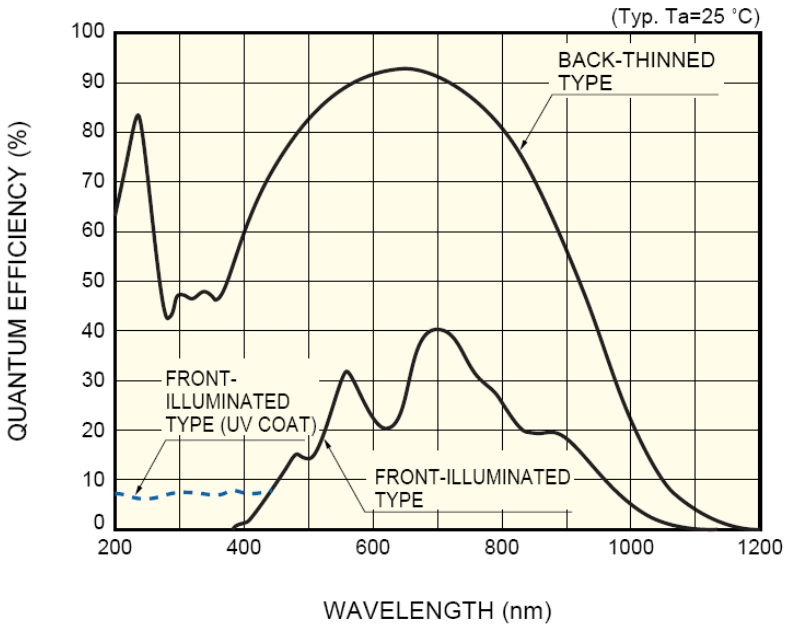


Figure 3.5 Spectral response function of front-illuminated CCD and Back-thinned CCD (back-illuminated) (reprinted from Ref. [210]).

From the discussions above, a dispersive Czerny-Turner spectrograph coupled with a NIR-enhanced BI-CCD was selected. This selected spectrometer model is QE65000 (Ocean Optics, Dunedin, FL) equipped with an NIR-enhanced charge-coupled device (CCD) detector (S7031-1006, 1024 × 58 with pixel sizes of 24.6 μm, QE > 90%, Hamamatsu). The CCD is cooled by thermoelectric (TE) cooling system to reduce the dark current. The reason is that CCD dark current, as one of the major noise sources for CCD based dispersive Raman spectroscopy, decreases as CCD working temperature decreases. Hence, CCD cooling can reduce the noise level and eventually improve the low detection limit of system and the SNR of Raman system. TE cooling can cool CCD efficiently down to be at -20 ºC by a heat sink and natural convection.

3.4 Raman Probe Design

3.4.1 History of Fiber-Optic Raman Probe

Raman probe is to deliver the light to sample and collect the Raman photons from sample. The probe determines the sampling location and volume of tissue. The

collection efficiency of Raman probe has a huge impact on the sensitivity of Raman spectroscopy system and SNR. It is also recognized as the dominant factor for *in vivo* Raman measurement on human body as remote control, compact size and flexibility are required for *in vivo* measurements on most of the human organs. Therefore, great efforts have been taken to develop and improve Raman probe.

In the past three decades, with the advance in fiber optics, fiber-optic Raman probe became prevailing and enabled *in vivo* Raman measurement on various human organs [67, 78, 99, 174, 211-216]. In the early 1980s, McCreery et al reported the first design of fiber-optic probe for remote Raman spectrum acquisition [217, 218]. However, the application of this probe on biological tissues is limited due to the interference from the fiber material. During the propagation of the excitation light in the excitation and collection fibers, the fiber materials will be induced to produce interfering spectral signal [219]. For example, fused-silica produces Raman signal which interferes with true tissue Raman signal; impurities and dopants emit intense fluorescence which adds short noise to the collected spectral signal; fiber jacket materials may make additional interference. Therefore, a variety of filtering designs were proposed to eliminate the interference from fiber materials. Myrick et al (1990) designed a fiber-optic probe coupled with a pair of band pass and long pass filters as shown in Fig. 3.6(a) [220]. Band pass filter serves to block the laser noise and optical signal from fiber material while transmitting excitation light; long pass filter is transparent to tissue signal while eliminating elastically scattered excitation laser light. However, the excitation and collection fibers are obliquely oriented, leading to non-compactness. Tanaka et al (1996) proposed a probe design to enhance the detection of tissue Raman signal by adding a compound parabolic concentrator (CPC) at the distal end of the probe (Fig. 3.6(b)). It achieved a 7-fold increase in collection efficiency compared to the probe

without CPC [221]. Owen et al (1995) reported a similar probe design with holographic grating and notch filters as shown in Fig. 3.6(c) [222]. The excitation light is collimated and then diffracted into the bottom light path by a transmission grating. Subsequently, excitation light is selected from the diffracted light by a spatial filter and notch filter, and then focused onto tissue by an objective. The backscattered tissue signal is collected by the same objective and filtered by another notch filter and eventually coupled into collection fiber by another objective.

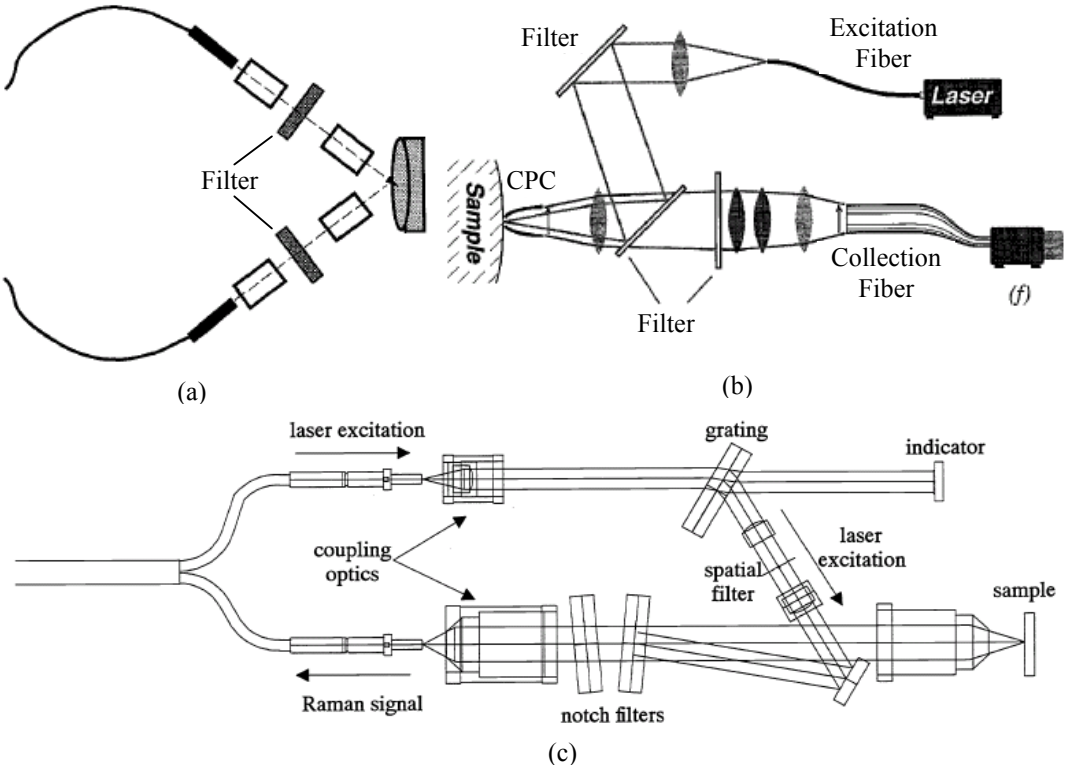
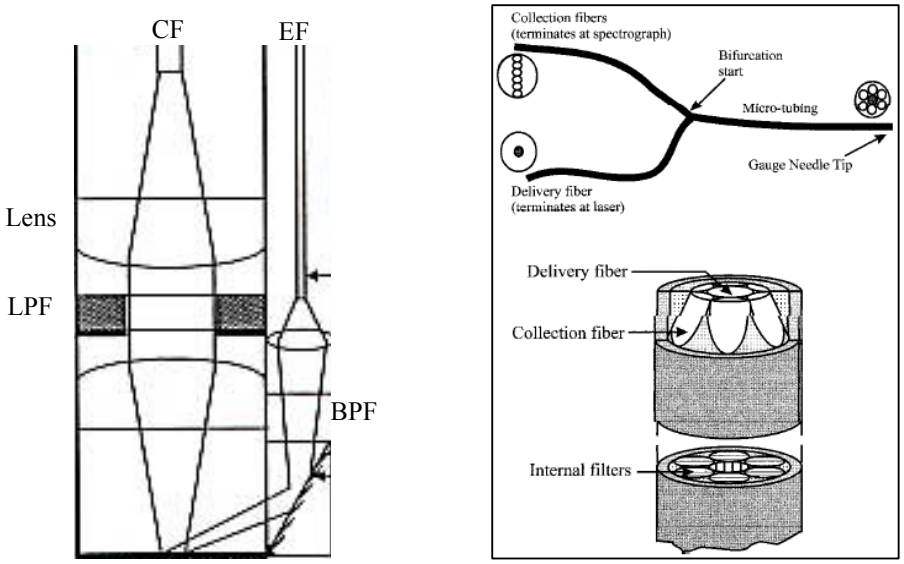


Figure 3.6 Schematic of three different fiber-optic Raman probe designs. CPC: compound parabolic concentrator ((a): adapted from Ref. [220]; (b): adapted from Ref. [221]; (c): reprinted from Ref. [222]).

Mahadevan-Jansen et al (1998) firstly designed a compact fiber-optic Raman probe which enables *in vivo* Raman measurement on human cervix as shown in Fig. 3.7(a) [99]. The excitation light from excitation fiber is focused by a 3-mm quartz lens and the focused light is directed onto the tissue surface by a golden mirror with oblique angle of 57°. The backscattered tissue signal is collected and focused into the

collection fiber by a pair of lens. A band pass filter and a long pass filter are inserted into the light delivery path and collection path, respectively. Raman spectra with clear Raman peaks can be acquired from cervix *in vivo* within 90 seconds. In the following year, Shim et al (1999) evaluated a bifurcate fiber-optic probe design with a central fiber for excitation light delivery and six surrounding fibers for tissue Raman signal collection as shown Fig. 3.7(b) [211]. Moreover, the collection fibers are beveled to increase the overlap of collection cone with excitation cone so as to improve the probe collection efficiency. This probe design also shows a good compatibility with endoscope for its compact size and flexibility and has proven to have the potential for acquiring Raman spectra *in vivo* from gastrointestinal tissues within 30 seconds [25]. Recently, it has been used for *in vivo* Raman measurement on cervix and only a short integration time of 5 seconds are required [101].



(a) (b)

Figure 3.7 Schematic of two fiber-optic Raman probe designs. CF: collection fiber; EF: excitation fiber; BPF: band pass filter; LPF: long pass filter ((a): adapted from Ref. [99]; (b): reprinted from Ref. [211]).

Motz et al (2004) introduced ball lens into fiber-optic Raman probe design (Fig. 3.8) to improve the collection efficiency [213]. This probe is constructed by attaching a ball

lens to a n-to-1 round bifurcate probe. A long pass filter tube and a short pass filter are incorporated between excitation/collection fibers and ball lens. With the use of this probe, high-quality Raman spectrum can be acquired within 1 second. Santos et al (2005) proposed a single fiber probe for measuring HW Raman spectrum since silica does not produce Raman and fluorescence interference with tissue Raman in HW region [174]. More recently, Matsuura et al (2007) developed a single fiber probe using hollow fiber with a ball lens installed at the distal end [216]. This single fiber probe allows tissue Raman measurement in both fingerprint and HW regions.

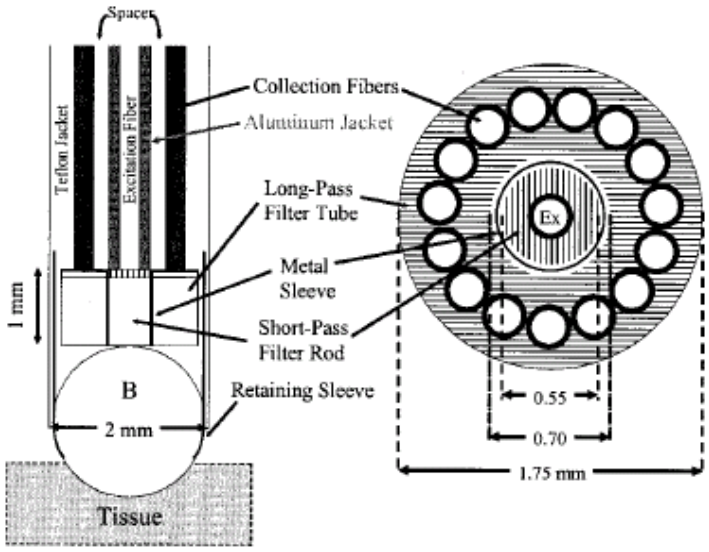


Figure 3.8 Schematic of fiber-optic Raman probe with a ball lens (reprinted from Ref. [213]).

In our early work, we developed a fiber-optic probe (Fig. 3.9) for *in vivo* rapid Raman measurement on skin [67]. A n-to-1 round fiber bundle adapter was specially designed and arrayed in parabolic line to correct the image aberration of spectrograph. This special design allows complete CCD vertical binning, leading to a 3.3- to 16-fold improvement in SNR. With this fiber-optic probe, good quality Raman spectrum can be acquired within 0.5 second. This probe has also proven its capability for *ex vivo* Raman measurement on human organs, including lung [51], larynx [45], nasopharynx [223], colon [28], stomach [70, 72]. In our recent work, we successfully developed an

endoscopic bifurcate fiber-optic Raman probe (32-to-1 round, outer diameter: 1.8 mm) which can be used to acquire Raman spectrum from stomach *in vivo* with 1 second or even shorter time through the biopsy channel of conventional endoscope under the guide of wide-field endoscopic imaging modalities.

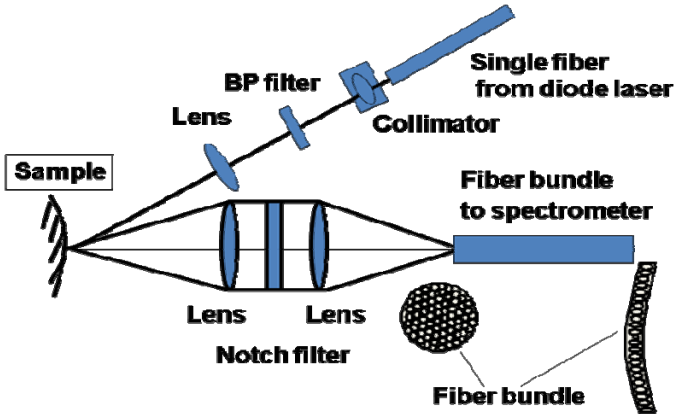


Figure 3.9 Schematic of a fiber-optic Raman probe (adapted from Ref. [67]).

3.4.2 Raman Probe Design

This thesis study is aimed at exploring the potential of NIR Raman spectroscopy for the *in vivo* diagnosis of cervical precancer. Note that cervix comprises two layers, including the superficial epithelium and the underlying stroma layer. Precancerous change usually arises in the basal layer and spreads towards the epithelium surface. The changes of tissue morphology or biochemical constituents associated with disease transformation may be depth-dependent [133]. For instance, the dysplasia-related changes (precancer) may be associated with the thickening of epithelial tissue (e.g., cervix), which results in an attenuation of the excitation light to penetrate into deeper areas of tissue and also the attenuation of Raman emission from deeper tissue regions (e.g., stroma) [190, 192]. On the other hand, the changes of other optical properties (e.g., absorption coefficient, scattering coefficient, anisotropic factor, refractive index) of tissue are also correlated with tissue physiologic/pathologic status, significantly affecting the overall Raman signal collection from biomedical tissue [224]. However,

the overall Raman signals acquired from the tissue surface are usually a mixture of Raman signal originating from different tissue depths, which consequently results in difficulties in analyzing Raman spectroscopic information emitted from tissue for tissue diagnosis and characterization [78, 211]. Hence, to better understand the origins of Raman signals collected from tissue surface for further improving the diagnosis of epithelial precancer (i.e., CIN), it is highly desirable to develop a depth-resolved Raman spectroscopic technique for facilitating the depth-selective Raman measurement of cervical epithelial tissues.

A number of fiber-optic probe designs have been reported for depth-resolved optical spectroscopic measurements, but most work are centered on fluorescence and reflectance spectroscopy for tissue diagnosis [191, 225-229]. The depth-resolved fiber probe designs can mainly be classified into two types: (1) single-fiber probe in which the same fiber is used for both light excitation and reflectance/fluorescence/Raman collection, and (2) multiple-fiber probe in which separate fibers are used for light delivery and reflectance/fluorescence/Raman collection [190, 224-227]. Although the single-fiber probe with a smaller aperture (i.e., core diameter) shows a high sensitivity of detecting spectral signal particularly emitted from the superficial layer of epithelial tissue, its collection efficiency is much reduced when the fiber diameter decreases. With multiple-fiber probe designs, varying the excitation-collection fiber separations or the oblique angles between the excitation-collection fibers permit the depth-discrimination of fluorescence/Raman signals from tissue [226, 227]. Although the multi-fibers probe design has shown a better depth-selectivity for spectroscopic acquisitions, the bulky probe design is unsuited for endoscope-based biomedical applications. In this work, we proposed a fiber-optic Raman probe design coupled with a ball lens for improving the collection efficiency of depth-resolved Raman

measurements, particularly for probing subsurface regions of epithelial tissue (e.g. cervix). The use of a single ball lens as both illumination and collection optics of spectroscopic probes allows a complete overlap of illumination and collection cone and has proven to be efficient for epithelial tissue measurement [213]. Moreover, the use of a ball-lens probe also shows some advantages, such as compactness and short working distance (probe-tissue contact configuration), which are in favor of *in vivo* tissue measurement.

Figure 3.10 depicts the schematic of ball lens Raman probe design. The Raman probe comprises two optical arms (one for laser light delivery; one for scattered tissue Raman collection) integrated with optical filtering modules. The 785-nm excitation light is coupled into the excitation arm of the Raman probe through a 200-μm fiber (NA=0.22), and delivered into the filtering module incorporated with a NIR lens for excitation light collimation and a narrow band-pass (BP) filter (centered at 785 nm, FWHM= ±2.5 nm) for removing fiber fluorescence and laser noise, and then focused onto the tissue through a NIR-coated sapphire ball lens (5 mm in diameter, refractive index n=1.76) mounted on the tip of the Raman probe. The backscattered tissue Raman photons are efficiently collected by the same ball lens and subsequently reflected back into the collection arm through a dichroic mirror and a reflection mirror. The backscattered light is further filtered by using an edge long-pass (LP) filter (cut off at 800 nm) to block the Rayleigh scattered laser light while allowing the Raman scattered light to pass into the spectrometer through a specially designed round to line fiber bundle adapter (28×50 μm, NA=0.22) for maximizing tissue Raman signals detection through vertical binning of the entire CCD [67]. All the optics of the Raman probe is sealed into the stainless steel sleeve (outer diameter of 8 mm) with polytetrafluoroethylene gasket for *in viv*o cervical tissue measurements. Each spectrum

was acquired within 1 s with light irradiance of 1.6 W/cm^2, which is less than the ANSI maximum permissible skin exposure limit for a 785-nm laser light [177]. The estimated temperature rise is about 0.3 ℃ after 1 minute of 785 nm laser irradiation with an incident power of 15 mW on a spot diameter of 0.2 mm during Raman measurement, which is far below the level to generate cytotoxicity in tissues and cells [230].

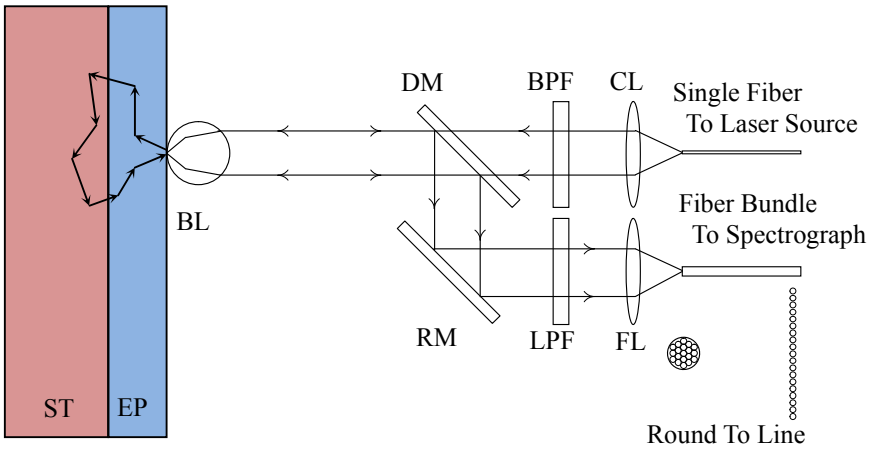


Figure 3.10 Schematic of the Raman probe design. CL: collimating Lens; FL: Focusing Lens; BPF: Band Pass Filter; LPF: Long Pass Filter; DM: Dichroic Mirror; RM: Reflection Mirror; BL: Ball Lens; EP: Epithelium; ST: Stroma.

3.4.3 Evaluation of Raman Probe Design by Monte Carlo Simulation

Monte Carlo (MC) simulations have been widely used to mimic light propagation in turbid media (e.g., biological tissue) [191, 229, 231], and also simulate the fluorescence generation and propagation in biological tissue [191, 229, 231-233]. With the development of MC simulation model to evaluate the fiber probe design coupled with a ball lens for depth-resolved fluorescence measurement on layered epithelial tissue [191, 229], we further modified the MC model to study Raman light generation and propagation in epithelial tissue as well as Raman signal collection with the use of ball lens fiber-optic Raman probe. The modified MC codes include the following modules: (1) excitation light propagation in ball lens and tissue; (2) Raman photons

generation and propagation in tissue; (3) Raman photons collection by the ball lens and propagation in the ball lens; (4) Raman photons collection by the focusing lens and subsequently by the collection fibers of the Raman probe.

In MC simulations, one hundred million excitation light photons are launched into the tissue through the ball lens. The photons are initialized as a uniform light beam with a beam diameter the same as the ball lens' diameter. When the excitation light photons are incident into the ball lens or the Raman photons pass through the ball lens, the direction of the refracted light is calculated using Snell's law. The photon weight is then updated as w←w(1-*R*), whereby the reflection coefficient, *R*, is calculated using Fresnel's formula:

$$R = \frac{1}{2}\left[\frac{\sin^2(\alpha_i - \alpha_t)}{\sin^2(\alpha_i + \alpha_t)} + \frac{\tan^2(\alpha_i - \alpha_t)}{\tan^2(\alpha_i + \alpha_t)}\right] \tag{3.1}$$

where α_i and α_t are the incident angle and the refracted angle, respectively. In this simulation, Raman generation is estimated with the probability (p) of the excitation light photons to be converted to Raman scattered light photons as follows [224, 234, 235]:

$$p = \left(\frac{\mu_s}{\mu_a + \mu_s}\right) \times RY \tag{3.2}$$

where μ_a is the Raman scattering coefficient for generating Raman photons, μ_s is the scattering coefficient, RY (<<1) is the Raman quantum yield. To determine the Raman photon generation at a particular scattering position, a random number between 0 and 1 is generated; if the probability yielded in Eq. (3.2) is larger than the random number, then Raman photon is produced; otherwise, no Raman photon is generated. In our MC programs, we presume that the radiation angle of Raman photons emitted from the Raman scatterers is isotropic [224, 236] and the emitted Raman photons will not be reabsorbed by the Raman scatterers for generating secondary emission of Raman

photons due to their extremely low intensity levels. The deflection angle (β) of each scattering in the medium is described by Henyey-Greenstein phase function as the following equation [231, 237]:

$$\cos\beta = \begin{cases} \dfrac{1}{2g}\left\{1+g^2-\left[\dfrac{1-g^2}{1-g+2g\xi}\right]^2\right\} & if \quad g>0 \\ 2\xi-1 & if \quad g=0 \end{cases} \tag{3.3}$$

where g is anisotropy factor and ξ is a random number between 0 and 1. The azimuthal angle ψ can be determined by 2πξ.

After escaping from tissue surface, Raman photons are collected by the ball lens within the collection cone angle [238]:

$$\theta = \sin^{-1}\left(\frac{2(n_b-1)}{n_b n_t}\right) \tag{3.4}$$

where θ is half-angle of the collection cone of the ball lens; n_b and n_t are the refractive indices of the ball lens and tissue, respectively. Tissue Raman photons which pass through the ball lens will be coupled into the collection fibers by a focusing lens (focal length, f=12 mm) of the Raman probe. Tissue Raman photons after ball lens are imaged into the collection fiber side by the focusing lens according to the following equation [239]:

$$\frac{1}{L'}+\frac{1}{L}=\frac{1}{f} \tag{3.5}$$

where L, L' and f are object distance, image distance, and focal length, respectively. The image position of Raman photons determines the light trace after focusing lens, which further determines the incident angle and position on the collection fiber surface plane. All those photons that reach the collection fibers within the collection cone angle of the collection fibers (NA=0.22) will be considered as the collected Raman photons by the Raman probe. The 2-D distributions of excitation light and the origins of the Raman photons collected from tissue will be given by MC simulations to

evaluate the performance of the fiber-optic Raman probe design for depth-resolved tissue Raman measurements. In addition, the collection efficiency of the Raman probe defined as the ratio of the collected tissue Raman photon numbers to the incident excitation light photon numbers is also estimated by using MC simulations.

A two-layer tissue model which mimics the epithelial tissues is used for MC simulations of the excitation light and Raman photons propagation in tissue. Table 3.3 gives the optical properties (Raman scattering coefficient, μ_a; scattering coefficient, μ_s; anisotropy factor, g; tissue refractive index, n_t; and tissue thickness, d) of the modeled tissue [240]. The wavelengths of excitation light and Raman scattered photon light are selected at 785 and 820 nm, respectively, for MC simulations [51, 67].

Table 3.2 Optical properties of the two-layer epithelial tissue model for MC simulations [191, 229, 240].

Tissue Layer	Wavelength (nm)	μ_a (cm^{-1})	μ_s (cm^{-1})	g	n_t	d (mm)
Epithelium	785 (excitation)	0.9	25	0.9	1.4	0.3
Stroma		0.45	165	0.9	1.4	3
Epithelium	820 (Raman)	0.8	24	0.9	1.4	0.3
Stroma		0.3	150	0.9	1.4	3

3.4.3.1 Effect of the Refractive Index of the Ball Lens on Tissue Raman Measurements

Figure 3.11 shows the MC simulation results of the 785-nm excitation light distribution in epithelial tissue for the fiber-optic Raman probe designs with different refractive indices of the ball lens and a fixed diameter of 3 mm. Note that the indices selected for simulation correspond to different optical materials (1.46-UV fused silica; 1.51-Boro-crown glass; 1.63-Dense flint glass; 1.76-Sapphire; 1.83-Lanthanum flint glass). Overall, the excitation light exhibited higher intensity over the whole tissue when increasing the refractive index of the ball lens. A 1.1- to 1.5- fold increase in

excitation intensity distribution in tissue was found for the use of ball lens at refractive indices of 1.51, 1.63, 1.76 and 1.83 as compared to the refractive index of 1.46. The corresponding excitation light distribution in the epithelium was approximately 1.1- to 2.7- fold stronger. This suggests that Raman probe with a larger refractive index of the ball lens achieve higher efficiency in focusing the excitation light photons into the tissue, especially in the epithelium layer of the tissue.

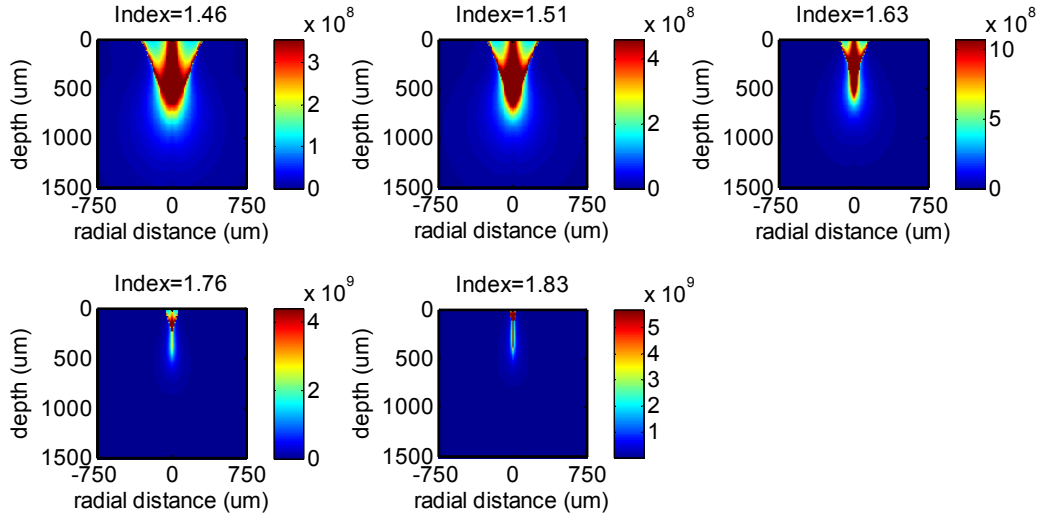


Figure 3.11 The 785-nm excitation light distributions along the tissue depth and radial directions in tissue using the Raman probe designs with different refractive indices of the ball lenses (*n*= 1.46, 1.51, 1.63, 1.76 and 1.83).

Figure 3.12 shows the 785-nm excitation light distributions along the tissue depth which are the sum of the excitation light distribution (Fig. 3.11) over the radial dimension at each tissue depth. Each depth profile was normalized to the integrated area under the profile for better comparison among different refractive indices of the ball lens in Raman probe. The excitation light underwent a shorter penetration into tissue when increasing the refractive index of the ball lens. For instance, an intensity maximum of the excitation light in epithelium layer was observed around the focal plane of the ball lens when using the refractive index of 1.76 or 1.83 of the ball lens (Fig. 3.12). This suggests that increasing the refractive index of the ball lens of the

Raman probe design make the excitation light to be more tightly focused in the shallower regions of tissue (i.e., the epithelium). The findings above can be explained by the enhanced focusing ability of the ball lens attributed to the increase of refractive index. The ball lens focal length can be estimated using the following equations [238]:

$$BFL = EFL - \frac{D}{2}; \quad EFL = \frac{n_b D}{4(n_b - 1)} \tag{3.6}$$

where BFL is back focal length, representing the distance from the ball lens front tip to the focal point; EFL is effective focal length, representing the distance from the ball lens center to the focal point; D and n_b is the diameter and the refractive index of the ball lens, respectively. According to Eq. (3.6), if keeping diameter D constant, EFL and BFL are inversely proportional to refractive index n_b. That is why the ball lens with a larger refractive index leads to a shorter light penetration depth in tissue.

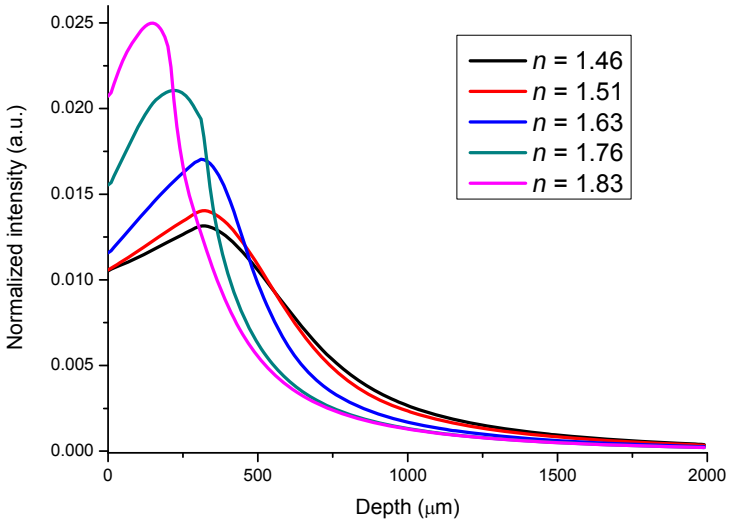


Figure 3.12 The 785-nm excitation light distributions along the tissue depth using the Raman probe designs with different refractive indices of the ball lenses (n=1.46, 1.51, 1.63, 1.76 and 1.83).

Figure 3.13 shows the calculated results of the origins of the Raman photons generated from different tissue locations of epithelial tissue. Similar to the excitation light distribution in tissue (Fig. 3.11), the Raman photons generated from different tissue depths were also dependent on the use of the refractive index of the ball lens in Raman

probe designs. With the increase of the refractive index of the ball lens, Raman photons radiation tended to be more strongly coming from the shallower regions of the epithelium layer (e.g., ~5 orders of magnitude stronger in Raman intensity using the ball lens's refractive index of 1.83 as compared to that using the ball lens's refractive index of 1.46) (Fig. 3.13). This is due to the fact that excitation light is more tightly focused in the shallower layer of the epithelium, and the ball lens with a larger refractive index has larger collection cone angle (Eq. (3.4)), leading to the significant improvements of Raman signal collections from the epithelium in tissue. The findings above imply that increasing the refractive index of the ball lens in Raman probe design can effectively improve the collection efficiency of tissue Raman photons from the tissue, especially from the subsurface regions of the epithelium layer in tissue.

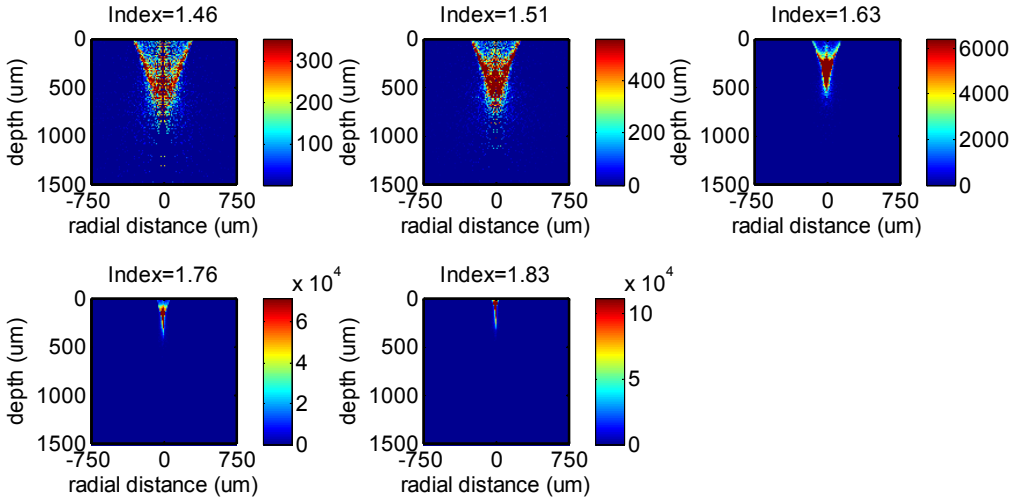


Figure 3.13 Distributions of the Raman photons collected from different tissue depths using the Raman probe designs with different refractive indices of the ball lens (n=1.46, 1.51, 1.63, 1.76 and 1.83).

Figure 3.14 shows the depth distribution of the origins of Raman photons collected from tissue for the use of different refractive indices of the ball lens in Raman probe designs. The depth distribution was created by integrating the 2-D distribution (Fig. 3.13) over the radial dimension at each tissue depth and was then normalized to the integrated area under the profile. It can be used to represent the relative contribution of

Raman photons from different tissue depths to the overall collected Raman photons. The relatively weak Raman photons were generated throughout different tissue depths when using the ball lens with refractive indices of less than 1.63 (Fig. 3.14); but stronger Raman photons generation tended to be arising from the epithelium layer in tissue when using the ball lens with larger refractive indices (e.g., n=1.76, 1.83). For instance, a maximum Raman radiation appeared in the shallower region (~80 μm) of the epithelium layer in tissue for the use of ball lens with refractive index of 1.83 (Fig. 3.14). This indicates that the Raman probe design with the use of high refractive index ball lens is in favor of detecting Raman signal emitted from subsurface layer of the epithelial tissue.

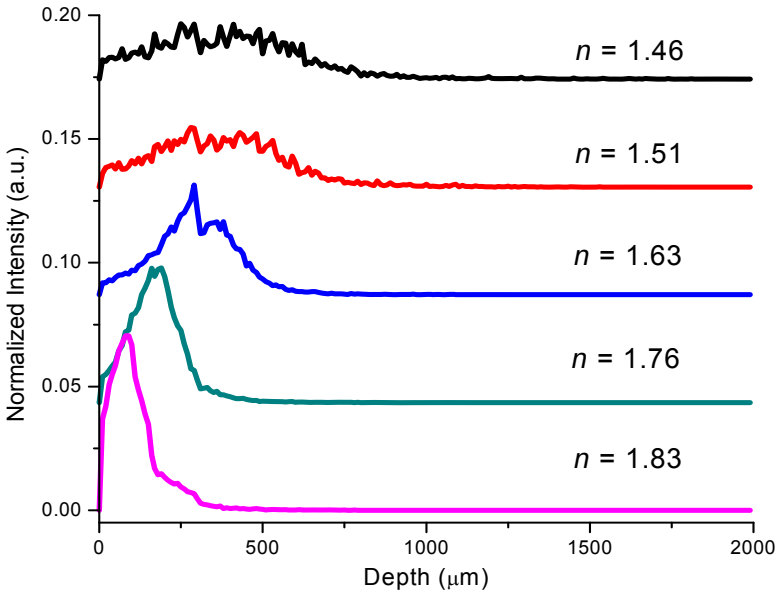


Figure 3.14 Depth-resolved Raman photons collected from different tissue depths using the Raman probe designs with different refractive indices of the ball lens (*n*=1.46, 1.51, 1.63, 1.76 and 1.83). Note that for comparison purpose, the depth-resolved Raman intensity profiles with different refractive indices of the ball lens have been vertically shifted to different intensity levels.

We also evaluated the collection efficiency of the Raman probe versus the refractive index of the ball lens (Fig. 3.15). The collection efficiency of the Raman probe increased monotonously when increasing the refractive index of 1.46 up to 1.83. All

the MC simulation results above confirm that Raman probe design with the use of a larger refractive index of the ball lens has the advantage of preferentially detecting the Raman signal from subsurface regions of the epithelium in tissue with high collection efficiency.

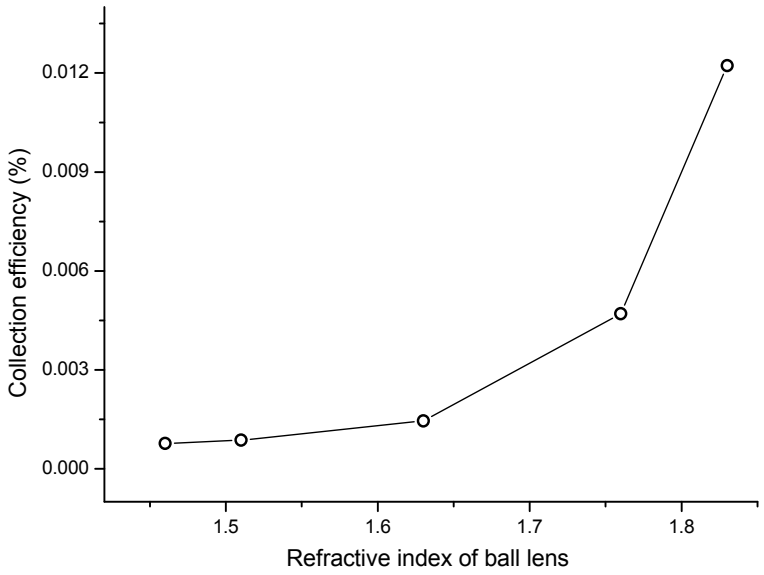


Figure 3.15 Collection efficiency of the Raman probe as a function of the refractive index of the ball lens.

3.4.3.2 Effect of the Ball Lens Diameter on Tissue Raman Measurements

Figure 3.16 shows the MC simulation results of the 785-nm excitation light distribution in epithelial tissue for the fiber-optic Raman probe designs with different diameters of the ball lens (refractive index of 1.83). With the use of smaller diameters of the ball lens in Raman probe design, the incident light was found to be more tightly focused into shallower regions of epithelial tissue with stronger incident light distribution as compared to the use of large diameters of the ball lens. This can be explained by the shortening of the ball lens focal length (Eq. (3.6)) [238] when using the ball lens with small diameters.

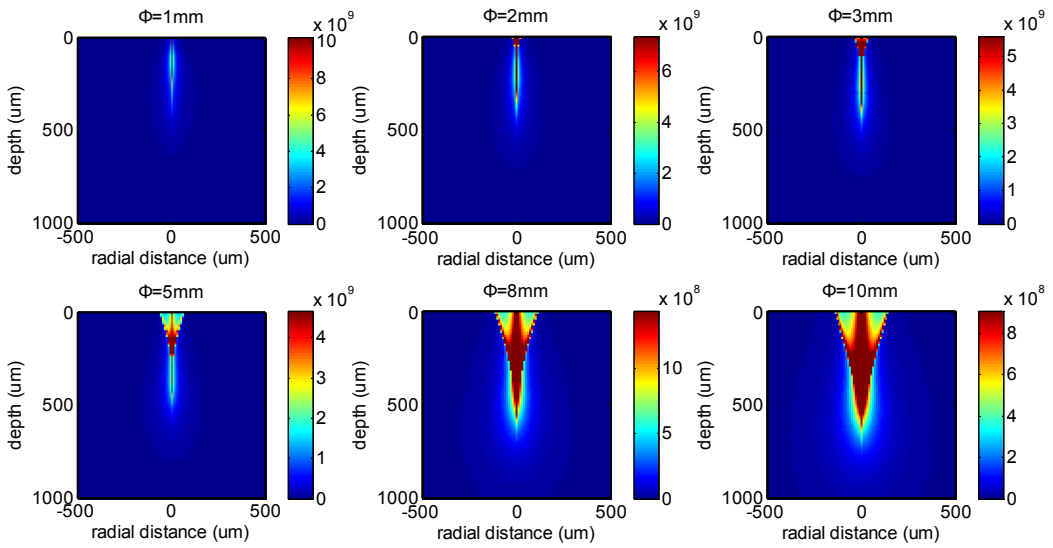


Figure 3.16 The 785-nm excitation light distributions in tissue using the Raman probe designs with different diameters of the ball lens (Φ=1, 2, 3, 5, 8 and 10 mm).

Figure 3.17 presents the depth distribution of the 785-nm excitation light in tissue calculated by integrating the 2-D intensity distribution (Fig. 3.16) over the radial dimension at each tissue depth. The excitation light from the Raman probe appeared to have a shorter penetration depth in tissue when using a smaller diameter of the ball lens. For instance, an incident light intensity maximum occurred at the subsurface region (~40 μm) of the epithelial tissue when the Raman probe used the ball lens with a diameter of 1 mm. Whereas the incident light tended to penetrate into deeper regions of the epithelial tissue when using the ball lens with larger diameters (e.g., 10 mm) (Fig. 3.17). Again, this implies that the Raman probe design with the use of smaller diameters of the ball lens can be more efficiently focusing the excitation light into the shallower regions of the epithelium layer in tissue.

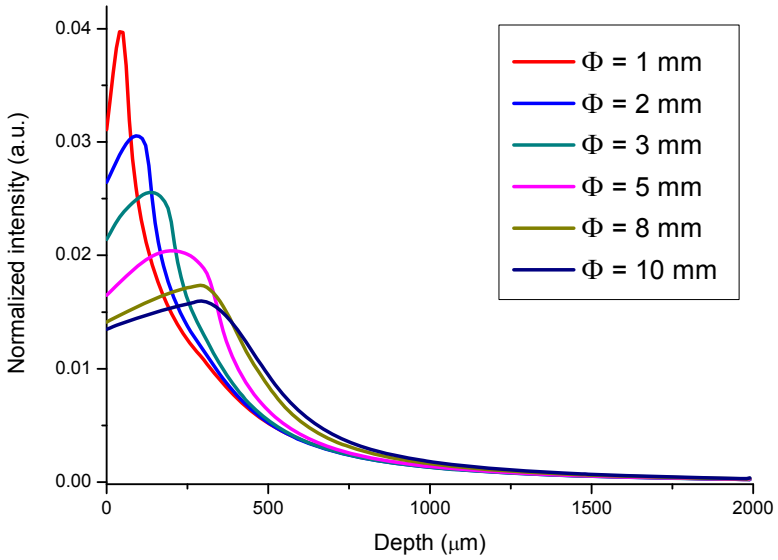


Figure 3.17 The 785-nm excitation light distribution along the tissue depth using the Raman probe designs with different diameters of the ball lens (Φ=1, 2, 3, 5, 8 and 10 mm).

Figure 3.18 depicts the 2-D distributions of the origins of the Raman photons collected from tissue using the Raman probes with different diameters of the ball lens. Similar to the corresponding excitation light distribution in tissue (Fig. 3.16), the tissue Raman photons collected with the use of a small diameter of the ball lens in Raman probe design appeared to be mostly originating from the shallower areas of the epithelium layer in tissue as compared to the Raman probe design with larger diameters of the ball lens. This can be more clearly displayed in the depth distribution of the origins of the Raman photons collected from tissue using the Raman probe designs with different diameters of the ball lens (Fig. 3.19). The relative contribution of the Raman photons from the epithelium became more dominant when reducing the ball lens' diameter in Raman probe design. For example, the Raman probe design with the use of 1-mm ball lens enabled a detection of ~95% of Raman signal arising from the epithelium, while only about 50% of Raman signal originating from the epithelium can be detected if using the 10 mm ball lens. Hence, the Raman probe design with a smaller diameter of ball lens can efficiently improve the detection of the Raman signal originating from the

epithelium layer in tissue.

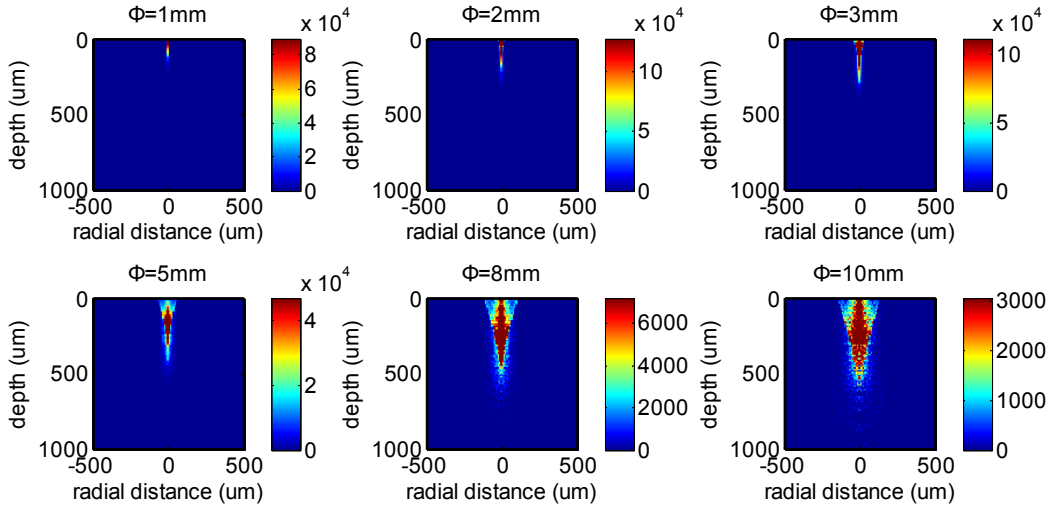


Figure 3.18 Distributions of the origins of the Raman photons collected from tissue using the Raman probe designs with different diameters of ball lenses (Φ=1, 2, 3, 5, 8 and 10 mm).

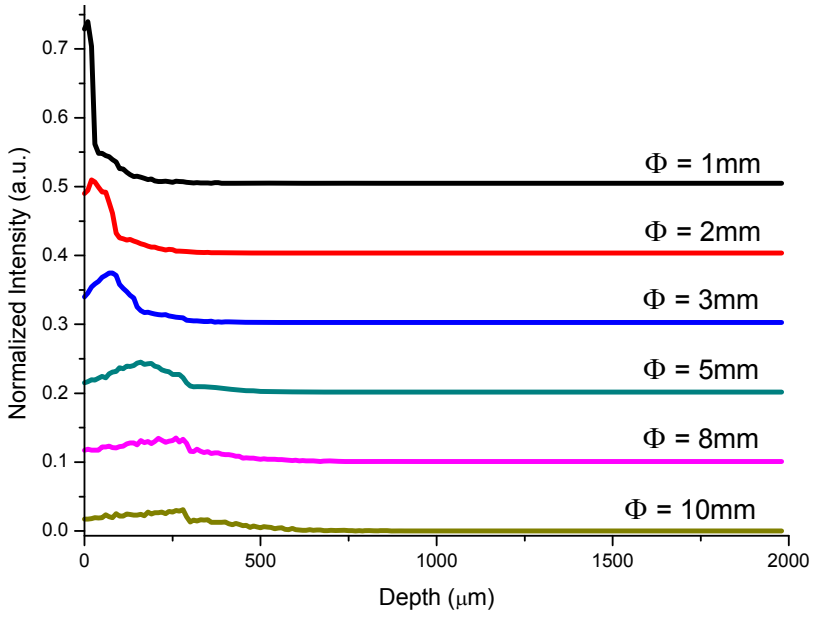


Figure 3.19 Depth distribution of the origins of the Raman photons collected from tissue using the Raman probe designs with different diameters of the ball lens (Φ=1, 2, 3, 5, 8 and 10 mm).

MC simulations were also carried out to exhaustedly search for all the possible combinations of the diameters and the refractive indices of the optimized Raman probe designs for improving depth-resolved Raman measurements. Figure 3.20 describes the depth-resolved measurement performance of the probe design as a function of the

combination of the ball lens' diameter and refractive index. The results show that the ball lens with a refractive index of 1.83 and diameter of 3 mm achieved the highest collection efficiency (Fig. 3.20(a)). With the use of this probe design, Raman photons originating in the epithelium of tissue covered 90% of the overall Raman photons collected from the surface of epithelial tissue as shown in Fig. 3.20(b), indicating a good depth-selectivity to the epithelium. Although the depth-selectivity can be further improved by reducing the diameter of the ball lens down to 2 mm, the collection efficiency will drop by about 30%. This implies that the Raman probe design with the use of a ball lens with the diameter of 3 mm and the refractive index of 1.83 is optimal for selectively detecting the Raman signal from the epithelial tissue. The above results indicate that Raman probe designs with a proper selection of the refractive index and diameter of the ball lens can improve both the collection efficiency and the ability to preferentially acquire the Raman photons emitted from the epithelium layer of epithelial tissue. In our clinic Raman system, Raman probe is configured with a sapphire ball lens at refractive index of 1.76 and diameter of 5 mm other than the optimal configuration suggested by the simulation results above. This is due to the fact that sapphire produces low interference signal and therefore is chosen as the ball lens material. The refractive index of sapphire is about 1.76 at 785 nm. Meanwhile, the reason we chose diameter of 5 mm is to increase the probe's sensitivity to the stroma layer which is rich in collagen, one of the major diagnostic biomarkers in cervix.

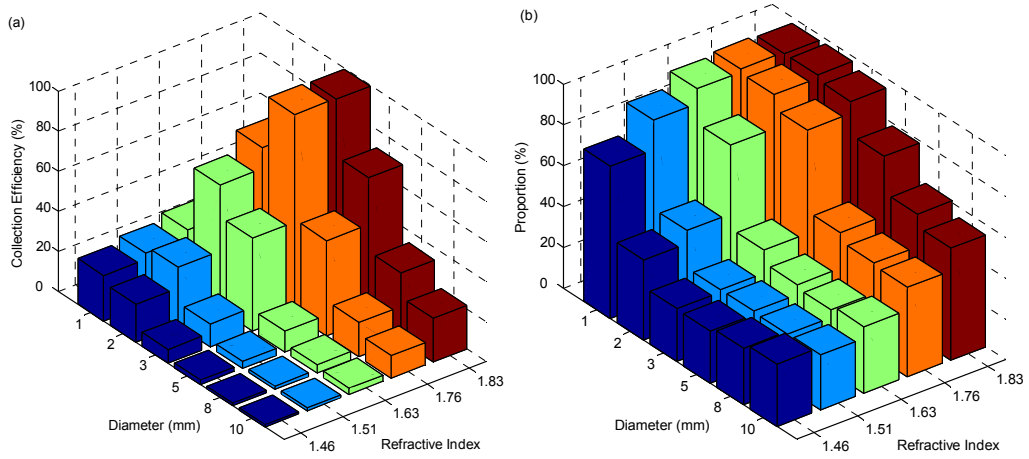


Figure 3.20 The simulated performance of the probe design based on various combinations of ball lens diameters (i.e., 1, 2, 3, 5, 8 and 10 mm) and refractive indices (i.e., 1.46, 1.51, 1.63, 1.76 and 1.83): (a) collection efficiency; (b) proportion of the Raman photons from the epithelium to the total collected Raman photons from the tissue surface.

3.4.3.3 Effect of the Probe-tissue Distance on Tissue Raman Measurements

In practical tissue measurements, the probe-to-tissue distance may play a significant role in depth-resolved tissue spectroscopic measurement [191, 225, 226, 229]. We also comprehensively studied the relationship of the collection efficiency of the ball lens Raman probe designs versus probe-tissue distances. Fig. 3.21(a) shows an example of the collection efficiency versus probe-tissue distances using the Raman probe with the ball lens' diameter of 3 mm and refractive index of 1.83. The maximum collection efficiency of the Raman probe occurred at the probe-to-tissue distance of ~300 μm. We also studied the effect of the probe-tissue distance on the ratio of the Raman signal collected from the epithelium layer to the overall Raman signal from the entire epithelial tissue in depth-resolved Raman measurements (Fig. 3.21(b)). It was found that when the probe-to-tissue distances were in the range of 275 to 660 μm, Raman photons arising from the tissue epithelium layer accounted for over 90% of the total Raman photons of the entire epithelial tissue. The above results suggest that the collection efficiency and the depth-selectivity of the Raman probe design can be optimized by setting a proper gap between the probe and the tissue.

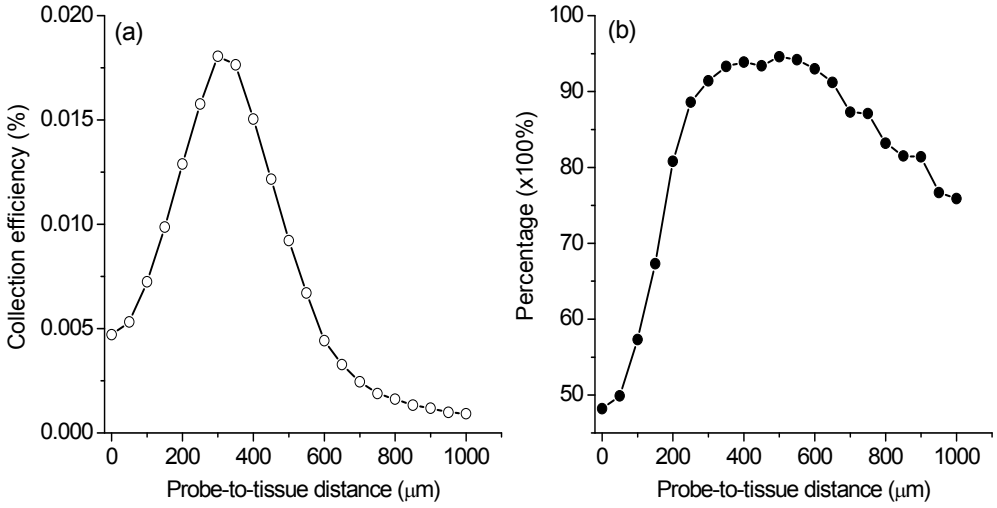


Figure 3.21 (a) Collection efficiency of the Raman probe as a function of probe-tissue distances; (b) Percentage of the Raman signal collected from the epithelium layer to the overall Raman signal from the entire epithelial tissue. The refractive index and the diameter of the ball lens are 1.83, and 3 mm, respectively, in Raman probe design.

3.4.4 Experimental Evaluation of Raman Probe Design

We also carried out tissue Raman measurements on a two-layer tissue phantom to evaluate the performance of a ball lens Raman probe design using a sapphire ball lens. The two-layer tissue phantom was constructed by overlaying different thicknesses of chicken muscle tissue on a chicken fat tissue: the thickness of fat tissue layer was fixed to be 5 mm, while the thickness of the muscle tissue layer varied from 0.3 to 3.9 mm. Figure 3.22 shows Raman spectra acquired from chicken muscle and fat tissue, as well as from the two-layer tissue phantoms with the muscle tissue thickness of 0.3, 1.2, 2.1, 3 and 3.9 mm, respectively. It was observed that chicken muscle and fat tissue showed distinctively different Raman spectral features. For instance, the Raman spectrum of muscle tissue showed distinct Raman peaks at 936 (C-C stretching of proteins), 1004 (C-C stretching of phenylalanine), 1209 (C-C_6H_5 stretching of tryptophan and phenylalanine), and 1339 cm^{-1} (CH_3CH_2 wagging of collagen). In contrast, the Raman spectrum of fat tissue exhibited four different unique peaks at 972 (Calcium-phosphate stretching of cholesterols/lipids), 1080 (C-C/C-O stretching of phospholipids), 1301

(CH_2 twisting and wagging modes of lipids, triglycerides (fatty acids)) and 1745 cm^{-1} (C=O stretching of ester (phospholipids)) [51, 190]. With the thickness of muscle tissue layer increasing from 0 to 3.9 mm in the two-layer tissue phantom, the Raman intensity at 1004 cm^{-1} for muscle tissue increased by 2.1- to 5.6-folds, whereas the Raman signal at 1745 cm^{-1} for fat tissue underwent a 28- to 1.6-fold reduction in intensity (Fig. 3.23). This implies that the Raman signal contribution of the surface muscle layer to the overall Raman signal increases as the muscle tissue thickness increases while the signal contribution of the deeper fat tissue layer reduces concomitantly. The results above indicate the potential of the ball lens Raman probe design for effective depth-selective Raman measurements in layered tissue.

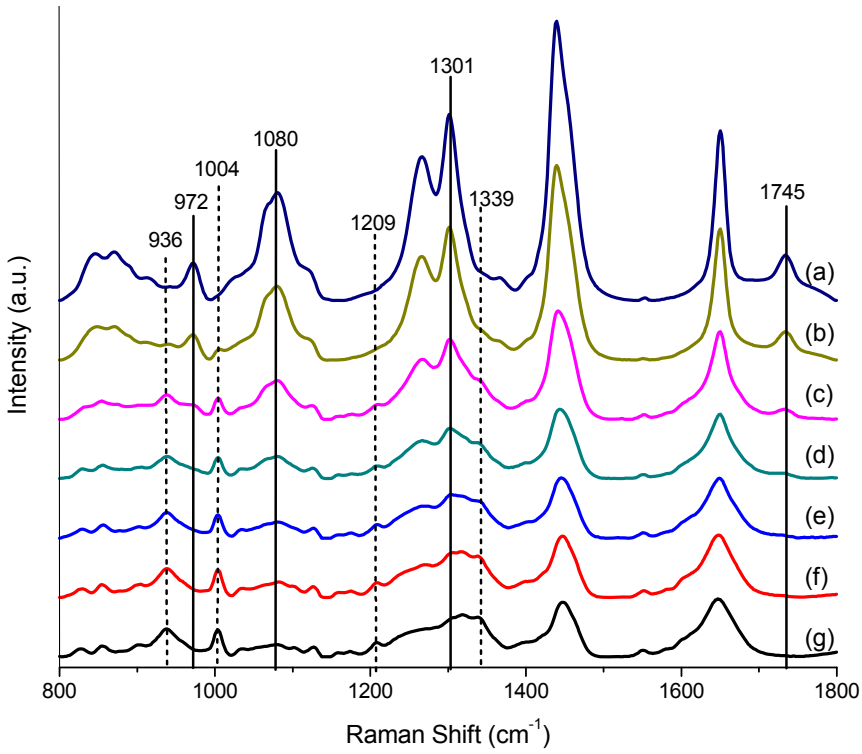


Figure 3.22 Raman spectra acquired from chicken muscle and fat tissue, as well as from the two-layer tissue phantoms with the muscle tissue thickness of 0.3, 1.2, 2.1, 3 and 3.9 mm, respectively. Spectra: (a) fat tissue; (b)-(f): two-layer tissue phantoms with the muscle tissue layer of thickness of 0.3, 1.2, 2.1, 3.0 and 3.9 mm, respectively, overlaying on a fat tissue layer (thickness of 5 mm); (g) muscle tissue. Note that all tissue Raman spectra are acquired with an integration of 1 s under the 785-nm excitation power of 1.5 W/cm^2. The dotted and solid vertical lines indicated in Raman spectra stand for the distinctive Raman peaks originating from the muscle and fat chicken tissue, respectively.

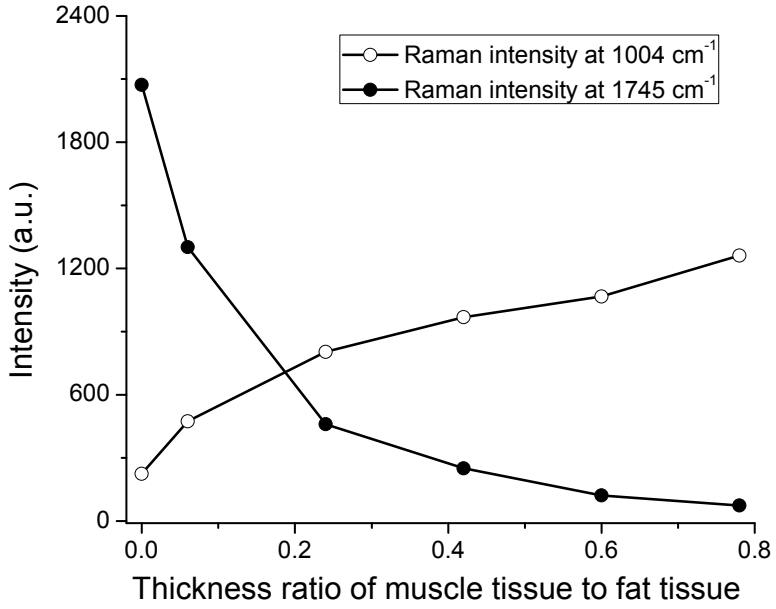


Figure 3.23 Raman peak intensities at 1004 and 1745 cm^{-1} as a function of thickness ratios of the muscle tissue layer to the fat tissue in a two-layer tissue phantom.

3.5 Data Acquisition Program

OOIBase32 (Ocean Optics, Dunedin, FL) is the official software for controlling Ocean Optics spectrometer and data acquisition. Besides, it also incorporates online dark-subtraction function as well as several basic spectral data process functions (e.g., smoothing by boxcar function, average consecutive spectra).

3.6 Conclusion

In summary, we successfully developed a rapid and portable Raman spectroscopy for *in vivo* Raman spectroscopic measurements on cervical tissues. A fiber-optic Raman probe was specially designed with the use of a ball lens for improving depth-resolved Raman measurements of epithelial tissue. We comprehensively evaluated the effects of optical configurations of the ball lens Raman probe design on depth-resolved Raman measurements using Monte Carlo simulations. The MC calculation results show that depth discriminations of Raman signals are achievable by changing either the refractive index or diameter of the ball lens, or the ball lens-tissue distances. The

Raman probe design by increasing the refractive index or reducing the diameter of the ball lens is in favor of probing Raman signals particularly arising from the shallower regions of the epithelium layer of tissue. Experimental evaluation of a ball lens probe design confirms its ability for efficient depth-resolved measurement on layered tissues. The efficacy of this Raman spectroscopy system for *in vivo* Raman measurement on cervical tissues will be discussed in chapter 4 and 5.

Chapter 4 High Wavenumber Raman Spectroscopy for *In Vivo* Detection of Cervical Dysplasia

The chapter is to present our work on applying NIR Raman spectroscopy in the high wavenumber (HW) region (2800~3700 cm^{-1}) for *in vivo* detection of cervical dysplasia. A rapid-acquisition NIR Raman spectroscopy system associated with a fiber-optic Raman probe was used for *in vivo* spectroscopic measurements at 785-nm light excitation. A total of 92 *in vivo* HW Raman spectra (46 normal, 46 dysplasia) were acquired from 46 patients with pap smear abnormalities in the cervix. Significant difference in Raman intensities of prominent Raman bands peaking at 2850 and 2885 cm^{-1} (CH_2 stretching of lipids), 2940 cm^{-1} (CH_3 stretching of proteins), and the broad Raman band of water peaking at 3400 cm^{-1} in the 3100~3700 cm^{-1} range were observed in normal and dysplasia cervical tissues. The diagnostic algorithms based on PCA and LDA together with the leave-one patient-out, cross-validation method on *in vivo* HW Raman spectra yielded a diagnostic sensitivity of 93.5% and specificity of 97.8% for dysplasia tissue identification. This study demonstrates for the first time that HW Raman spectroscopy has the potential for the noninvasive, rapid, *in vivo* diagnosis and detection of precancer in the cervix.

4.1 Introduction

NIR Raman spectroscopy has been applied for optical diagnosis of cervical dysplasia and Raman spectral differences between normal and dysplasia cervical tissues have been observed [98, 100], demonstrating the diagnostic ability of NIR Raman spectroscopy. To date, most of cervical NIR Raman studies are centered on the fingerprint region (i.e., 800~1800 cm^{-1}) owing to the great wealth of biochemical information contained in this spectral region for tissue characterization [51, 78, 100, 189]. However, the strong autofluorescence background and silica Raman signal arising from fiber optic Raman probe also fall into the fingerprint region which severely interfere the detection of inherently weak tissue Raman signal [98, 175]. This makes the light filtering modules in Raman probe design to be complicated and bulky, and unsuitable for *in vivo* biomedical applications. To tackle this problem, Puppels and co-coworkers [8, 174] proposed to use Raman spectroscopy in the so-called high wavenumber (HW) region (i.e., 2400~3800 cm^{-1}) for tissue diagnosis and characterization. The main advantages of using HW Raman spectroscopy are the significant reduction of fluorescence/Raman background from optical fibers [8], more intense tissue Raman signals generated compared to the fingerprint region [241], as well as the possibility of an unfiltered, simplified single fiber Raman probe design for facilitating *in vivo* Raman clinical procedures [242]. For instance, HW Raman studies [60] showed that basal cell carcinoma could be distinguished with a 100% prediction accuracy, suggesting that diagnostic information extracted from the HW Raman spectral region could compete with those extracted from the fingerprint region. To date, the clinical potential of HW Raman spectroscopy for identification of cervical dysplasia (i.e., precancer) has yet been reported in literature. Therefore, in this study, we aimed to evaluate the feasibility of utilizing HW Raman spectroscopy for *in vivo*

diagnosis of cervical precancer. Multivariate statistical techniques, including PCA and LDA, were employed to develop diagnostic algorithms for differentiation between normal and precancer cervical tissue. Receiver operating characteristic (ROC) testing was also conducted to further evaluate the performance of PCA-LDA algorithms on *in vivo* HW Raman spectroscopy for cervical precancer diagnosis.

4.2 Materials and Methods

4.2.1 Raman Instrumentation

The Raman spectroscopy used in this study has been described in detail in Chapter 3.

4.2.2 Patients

A total of 46 women (mean age of 42.6 years old) who underwent colposcopic examinations for abnormal pap smears were recruited through colposcopy clinics at the National University Hospital (NUH) of Singapore. All patients preoperatively signed an informed consent permitting the *in vivo* spectroscopic measurements on the cervix. This study was approved by the Ethics Committee of the National Healthcare Group of Singapore. Prior to Raman spectroscopic measurements, a complete routine colposcopic examination was performed on the patients by experienced colposcopists. A 5% acetic acid was applied to the cervix for a couple of minutes to allow the inspection of acetic acid whitening changes of epithelium which was then graded according to the degree of acetic acid-whiteness, the shape and margin of lesion area and the vascular pattern. Evaluation of Raman spectra acquired before and after the application of acetic acid indicates that acetic acid does not lead to any significant change on the tissue Raman spectra. *In vivo* HW Raman spectra were acquired from the suspicious lesion sites and the surrounding normal sites for each patient through

gently placing the ball-lens Raman probe on the cervix. The measured abnormal sites were biopsied and then submitted to pathology laboratories of NUH for histopathologic examination to confirm the colposcopic diagnosis. The measured normal sites are normal mature squamous epithelium without showing acetowhite. A total of 92 *in vivo* HW Raman spectra were acquired from 46 patients with cervical abnormalities, in which 46 were from normal sites, while 46 from cervical dysplasia lesions, including 10 CIN 1, 9 CIN 2 and 27 CIN 3.

4.2.3 Data Preprocessing

The raw HW Raman spectra (2800~3700 cm^{-1}) measured from *in vivo* cervical tissue represented a composition of Raman signal, autofluorescence background and noise. Thus, the raw spectra were preprocessed by a first-order Savitzky-Golay filter (window width of 3 pixels, which corresponded to the system spectral resolution) to reduce noise [178]. A first-order polynomial was found to be optimal for fitting the autofluorescence background in the noise-reduced spectrum, and this polynomial was then subtracted from the raw spectrum to yield the tissue HW Raman spectrum alone. Each of background-subtracted HW Raman spectra was also normalized to the integrated area under the curve from 2800 to 3700 cm^{-1}, enabling a better comparison of the spectral shapes and relative band intensities among the different cervical tissues.

4.2.4 Multivariate Statistical Analysis

PCA-LDA was employed for developing diagnostic algorithms and has been described in detail in the section **Multivariate Statistical Analysis** in Chapter 2.

4.3 Results

4.3.1 Spectral Feature Analysis

Figure 4.1(a) shows the comparison of mean *in vivo* HW Raman spectra ± 1 standard deviation (SD) of normal (n=46) and dysplasia (n=46) cervical tissues. Prominent Raman bands, such as 2850 and 2885 cm^{-1} (CH_2 stretching vibrations of lipids), 2940 cm^{-1} (CH_3 stretching vibrations of proteins), and the broad Raman band of water (OH stretching vibration band peaking at 3400 cm^{-1} in the 3100~3700 cm^{-1}) [8, 174, 241], were clearly observed in both normal and dysplasia cervical tissues. Overall, precancer cervical tissue showed significantly lower intensities for the Raman bands in the 2800~3000 cm^{-1} region (paired two-sided Student's t-test, $p<0.001$, n=46), while being higher for OH bands in the 3100~3700 cm^{-1} (paired two-sided Student's t-test, $p<0.001$, n=46) compared to normal tissue as shown in the difference spectrum in Fig. 4.1(b). Nevertheless, there were significant variations and overlapping in intensities of *in vivo* HW Raman spectra of normal and precancer tissue among inter-subjects (Fig. 4.1). Thus, tissue diagnosis simply based on spectral intensities alone was precluded. In addition, it is also difficult to standardize the absolute intensities of spectral measurements in clinical settings. Therefore, in this study we focused on analyzing spectral feature differences following a standardization/normalization data pre-processing procedure.

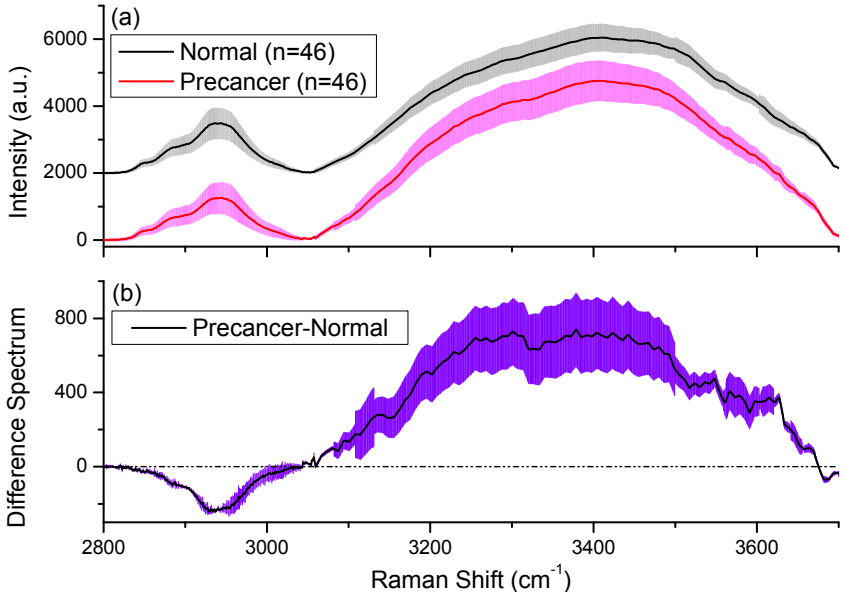


Figure 4.1 (a) Comparison of mean *in vivo* HW Raman spectra±1SD of normal (n=46) and precancer (n=46) cervical tissue. (b) Difference spectrum±1SD difference between precancer (n=46) and normal cervical tissue (n=46). Note that the mean *in vivo* HW Raman spectrum of normal tissue was shifted vertically for better visualization (Fig. 4.1(a)); the shaded areas indicate the respective standard deviations.

4.3.2 PCA-LDA and ROC Analysis

We employed the multivariate statistical methods (e.g., PCA, LDA) together with paired two-sided Student's *t*-test by incorporating the entire HW Raman spectra to determine the diagnostically significant Raman features for tissue diagnosis and classification. Figure 4.2 shows the first five PCs calculated from PCA on all HW Raman spectra. The first PC accounted for the largest variance (e.g. 49.6% of the total variance), whereas the successive PCs described the spectral features that contributed progressively smaller variances. Some PC features, such as peaks, troughs, spectral shapes, were similar to those of tissue HW Raman spectra (Fig. 4.1). The paired two-sided Student's *t*-test on the obtained PCs showed that there were three diagnostically significant PCs (i.e., PC1, PC4, and PC9) ($p<0.05$) for discriminating dysplasia tissue from normal tissue.

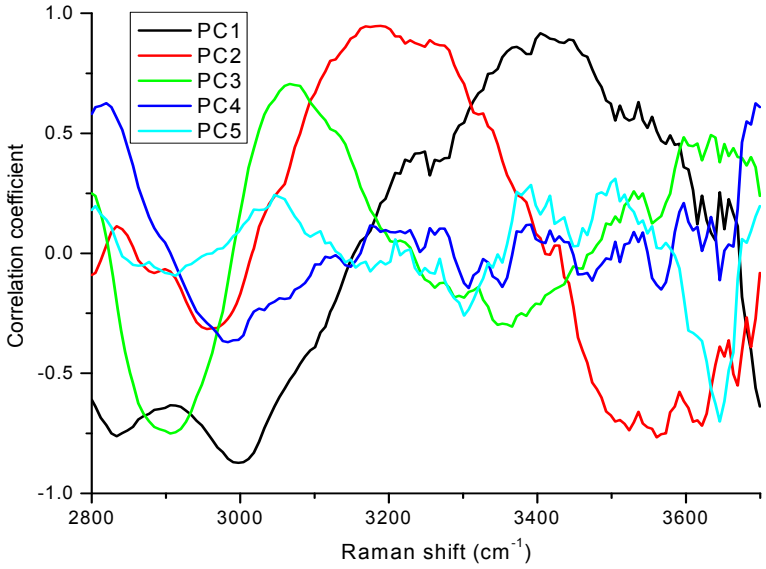


Figure 4.2 The first five principal components (PCs) accounting for about 88% of the total variance calculated from *in vivo* HW Raman spectra of cervical tissue (PC1-49.6%, PC2-21.7%, PC3-10.9%, PC4-4.7%, and PC5-1.6%).

Figure 4.3 shows the correlations between the diagnostically significant PC scores for normal and dysplastic cervical tissue, illustrating the utility of significant PCs for classification of Raman spectra between normal and precancer cervical tissues. Normal and dysplasia tissues can be largely clustered into two separate groups based on different combinations of significant PCs: (a) PC1 *vs* PC4, (b) PC1 *vs* PC9, and (c) PC4 *vs* PC9. The corresponding separation lines (i.e., diagnostic algorithms) classify dysplasia from normal tissue with the sensitivity of 63.0% (29/46), 89.1% (41/46) and 73.9% (34/46); specificity of 87.0% (40/46), 84.8% (39/46) and 87.0% (40/46), respectively. These results show that selection of different combinations of significant PCs will lead to different levels of accuracy for tissue classification.

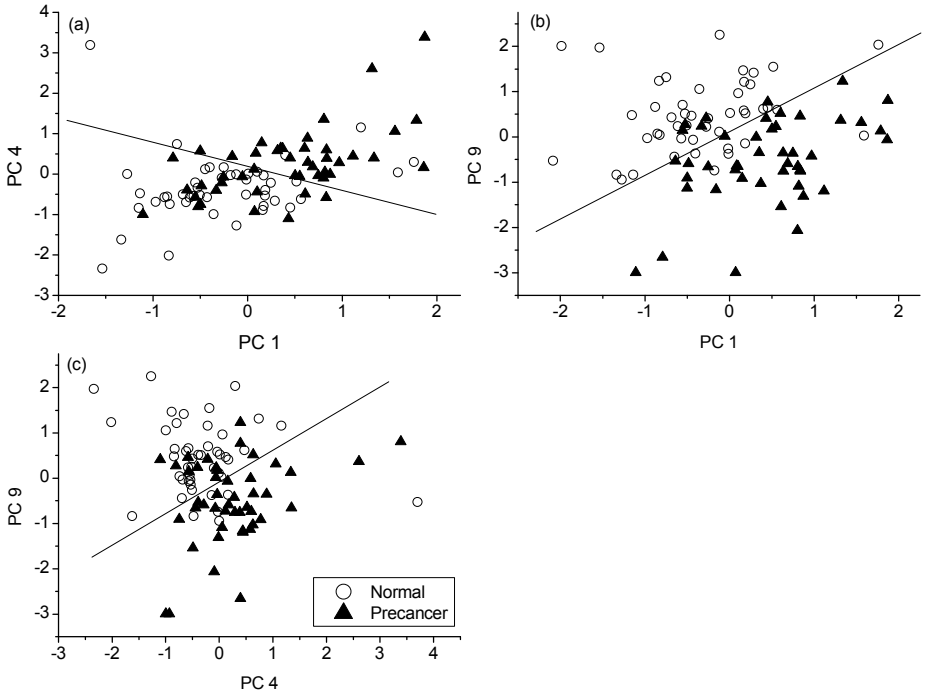


Figure 4.3 Scatter plots of the diagnostically significant PCs derived from *in vivo* HW Raman spectra of normal and precancer cervical tissue: (a) PC1 *vs* PC4; (b) PC1 *vs* PC9; (c) PC4 *vs* PC9. The dotted lines (PC4 = −0.57 PC1 + 0.19; PC9 = 0.96 PC1 + 0.12; PC9 = 0.62 PC4 − 0.08) as diagnostic algorithms classify precancer from normal with sensitivities of 63.0% (29/46), 89.1% (41/46) and 73.9% (34/46); specificities of 87.0% (40/46), 84.8% (39/46) and 87.0% (40/46), respectively. circle (○): Normal; triangle (▲): Precancer.

To further improve tissue diagnosis, all the three diagnostically significant PCs were loaded into the LDA model for generating effective diagnostic algorithms for tissue classification. Figure 4.4 shows the posterior probability of classification results based on PCA-LDA technique together with leave-one patient-out, cross-validation method. The PCA-LDA diagnostic algorithms yielded the diagnostic sensitivity of 93.5% (43/46) and specificity 97.8% (45/46) for distinguishing dysplasia from normal cervical tissue. To evaluate the performance of the PCA-LDA-based diagnostic algorithms derived from all the significant PCs of tissue HW Raman dataset, the receiver operating characteristic (ROC) curve (Fig. 4.5) was generated from the scatter plot in Fig. 4.4 at different threshold levels to determine the correct or incorrect classification of cervical tissues. The integration area under the ROC curve is 0.98, demonstrating the efficacy of PCA-LDA diagnostic algorithms developed that utilize

the entire spectral features of HW Raman spectroscopy for *in vivo* diagnosis of cervical precancer.

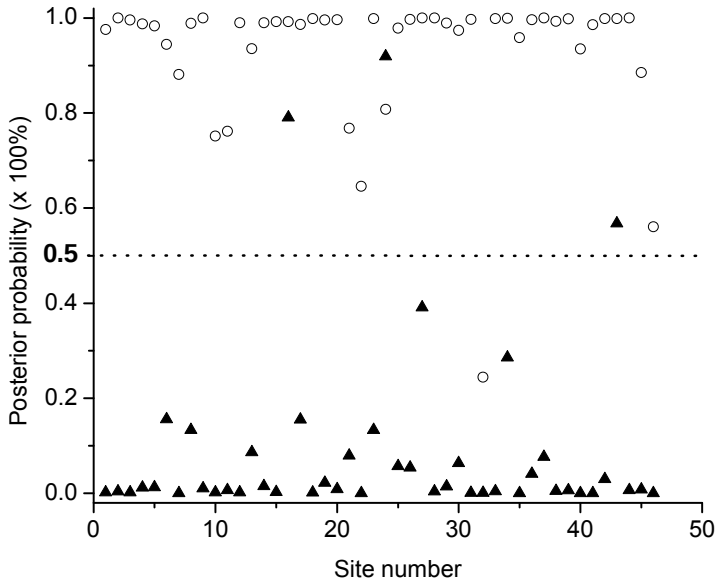


Figure 4.4 Scatter plot of the posterior probability of belonging to the normal and precancer cervical tissues using the PCA-LDA technique together with leave-one patient-out, cross-validation method. The separate line yields a diagnostic sensitivity of 93.5% (43/46) and specificity 97.8% (45/46), for identifying precancer from normal cervical tissue. circle (○): Normal; triangle (▲): Precancer.

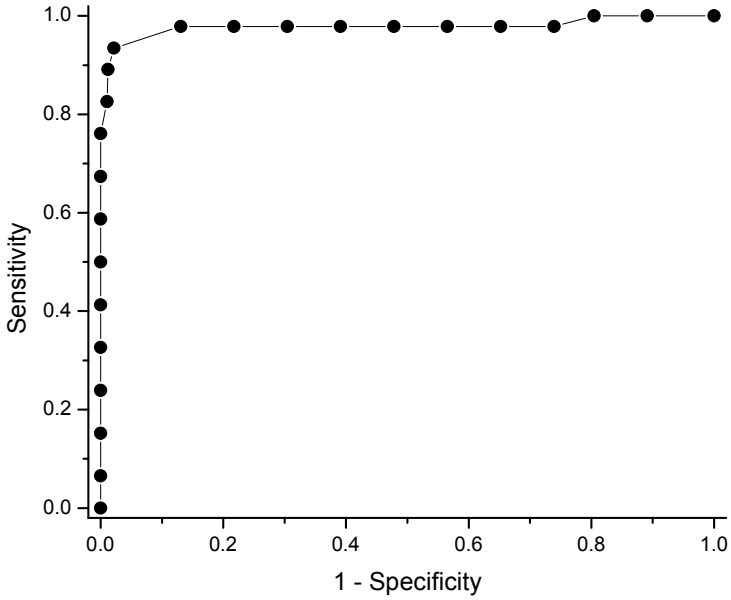


Figure 4.5 Receiver operating characteristic (ROC) curve of discrimination results for *in vivo* HW Raman spectra of cervical tissue using PCA-LDA algorithms together with leave-one patient-out, cross-validation method. The integration area under the ROC curve is 0.98, illustrating the efficacy of PCA-LDA algorithms for tissue classification.

4.4 Discussion

Raman spectroscopy has proven to be capable of probing biochemical structures and compositions in tissues associated with tissue malignant transformation. Likewise, Fourier transform infrared spectroscopy (FTIR), as the other vibrational spectroscopy technique, can also been used for tissue differentiation. Particularly for human cervix, FTIR spectroscopy and imaging technique has been successfully applied to the identification of the malignant cervical cells and tissues [155, 157, 170]. In principle, Raman and FTIR spectroscopy are complementary to each other due to the different selection rules for molecular vibrational modes [4], and have been recently demonstrated as complementary techniques for tissue identification [243]. However, FTIR is limited to *in vitro* spectroscopic measurement on human tissue. In contrast, Raman spectroscopy has shown the potential to be a clinically useful tool for precancer and cancer diagnosis in humans; however, due to the weak signal, an efficient and quick detection of tissue Raman signals is important for *in vivo* Raman clinical applications. In this work, we developed a rapid and portable NIR Raman spectroscopy system coupled with a ball-lens fiber-optic Raman probe [191, 229] to provide a high degree of collection efficiency for real-time *in vivo* cervical tissue measurements. Compared to the conventional fiber-optic Raman probes with a relatively small collection angle (of ~13 degrees; NA=0.22) [25], our Raman probe coupled with an NIR-coated sapphire ball lens (diameter of 5 mm; refractive index of 1.77) offers a large collection angle (of up to 55 degrees) to maximize the collection of back scattered Raman photons from tissue. In addition, the unique round-to-line fiber bundle adapter connected to the Raman probe optimizes the coupling of Raman photons collected into the spectrometer for further improving tissue Raman acquisitions [67]. As a result, good quality (SNR of up to 30 at 1 s integration times)

HW Raman spectra in the range of 2800~3700 cm^{-1} can be acquired from cervical tissue *in vivo* using the rapid Raman spectroscopy system developed.

We applied this rapid HW Raman spectroscopy for the first time for the *in vivo* diagnosis and detection of cervical precancer during clinical colposcopic examination. The overall intensity of Raman signals of cervical dysplasia in the region (2800~3050 cm^{-1}) involving proteins (e.g., 2940 cm^{-1}) and lipids (e.g., 2850 and 2885 cm^{-1}) were significantly lower than those in normal cervical tissue. Our MC simulations based on the two-layer (i.e., epithelium, stroma) tissue model according to the cervical tissue optical properties [192, 240] show that the 785-nm laser light from our ball-lens Raman probe can penetrate into the tissue depth of 690 μm in normal cervical tissue, while 630 μm in dysplastic tissue, thereby interrogating with cervical tissue volumes comprising both epithelium and stroma tissue in the cervix. Due to the changes of tissue optical properties (e.g., thickening of epithelium) and morphologies (e.g., higher cellular density resulting from the increased nucleus to cytoplasm ratios) associated with dysplastic transformation [72, 191], these effects significantly attenuate the excitation light penetration and also obscure the tissue Raman photons emitted from the underlying stroma in precancer tissue as compared to normal cervical tissue. Further MC simulations indicate that the overall Raman signals in the 2800~3050 cm^{-1} emitted from normal tissue are 1.15-fold stronger than dysplastic cervical tissue, which is in good agreement with the observation in our HW Raman cervical tissue measurements (Fig. 4.1). We have also studied the differences of the two respective tissue layers (i.e., epithelium, and stroma) in contributions to the HW Raman signals between normal and dysplastic cervical tissues. The MC results show that the stroma tissue (i.e., collagen) in dysplasia cervical tissue contribute only about 10% to the overall HW Raman signal in the 2800~3050 cm^{-1} as compared to the 50% of

contributions from normal stroma tissue. This also explains the significant reduction of collagen Raman signal (at 2940 cm^{-1}) observed in dysplasia tissue (Fig. 4.1). Similar results in the finding of much reduced collagen signals associated with cervical dysplastic changes have also been reported using fluorescence spectroscopy and Raman spectroscopy in the fingerprint region [100, 102, 134, 138]. On the other hand, we have also observed that the water vibrational signal in the 3100~3700 cm^{-1} is significant higher in dysplastic cervical tissue (Fig. 4.1), indicating the increase of water contents in percentage relative to all the Raman-active components in dysplastic cervical tissue. The increase of water concentrations associated with tumor alterations in other organs have also been reported using NIR spectroscopy techniques [244-246]. This phenomenon is probably due to the increase of certain types of aquaporins (AQPs) at dysplastic cervical cells which could increase the plasma membrane osmotic water permeability of up to 10-fold, thereby facilitating the influx of water molecules into the intracellular space in dysplasia tissue [247, 248]. Therefore, the distinctive spectral differences observed in the *in vivo* HW Raman spectra of normal and dysplastic cervical tissue in this study suggest the diagnostic potential of HW Raman spectroscopy for noninvasively distinguishing dysplastic tissue from normal cervical tissue.

To develop the diagnostic algorithms, empirical method, intensity ratio of specific bands [51], was employed prior to PCA-LDA analysis. The Raman band ratio of 2850 to 2940 cm^{-1}, ascribed to lipid/protein, was significantly higher in precancer compared to normal tissue. The result is in an agreement with Raman bands ratio finding in fingerprint region, higher ratio of 1445/1657 cm^{-1} (lipid/protein) in precancer [5]. Besides the reason of less contribution of collagen as discussed above, the higher ratio may also be attributed to the increased content of lipid relative to the total Raman

scatterers in tissue. This ratio of lipid/protein achieved a sensitivity of 78.3%, specificity of 89.1% and accuracy of 83.7% in differentiating dysplasia from normal tissues, confirming the lipid/protein ratio as a biomarker for cervical dysplasia. However, the ratio method only utilized a limited number of band intensities while neglecting most of Raman spectral features [70]. Moreover, biological tissue is very complex, it is thereby likely that there are many biochemical species influencing diseases concurrently [70, 244]. Therefore, in this work, the multivariate statistical analysis (e.g., PCA and LDA) was implemented by utilizing the entire HW Raman spectrum (2800~3700 cm^{-1}) to determine the diagnostically significant spectral features for dysplastic cervical tissue diagnosis and classification.

With the limited HW Raman spectral datasets we acquired to date, PCA-LDA together with leave-one patient-out, cross-validation technique was applied to the HW Raman spectra for dysplastic tissue identification in an unbiased manner. The cross-validated diagnostic sensitivity of 93.5%, specificity 97.8% and accuracy of 95.7% achieved for distinguishing dysplasia from normal cervical tissue confirm the diagnostic potential of HW NIR Raman spectroscopy for cervical precancer diagnosis. To further verify if the utilization of the entire HW spectral region (i.e., 2800~3700 cm^{-1}) is robust for providing good tissue classification accuracy, we have also evaluated the diagnostic abilities based on the two Raman sub-regions, i.e., 2800~3050 cm^{-1} (containing proteins and lipids biomolecular constituents), and 3100~3700 cm^{-1} (representing water molecule) for cervical precancer diagnosis. PCA-LDA analysis yielded a diagnostic sensitivity of 89.1%, specificity of 89.1% and accuracy of 89.1% in the spectral region of 2800~3050 cm^{-1}; and sensitivity of 84.8%, specificity of 89.1% and accuracy of 87% in the region of 3100~3700 cm^{-1}, respectively, for cervix dysplasia identification. The highest diagnostic accuracy (~96%) achieved by employing the entire HW

spectral region from 2800~3700 cm^{-1} compared to the truncated HW Raman sub-regions (i.e., 2800~3050 cm^{-1} and 3100~3700 cm^{-1}) further reinforces that the unique combination of different Raman biomolecular signals ranging from water, lipids to proteins (i.e., 2800~3700 cm^{-1}) can provide good differentiation between normal and dysplastic cervical tissue. The overall diagnostic accuracy rate of 96% achieved in the 2800~3700 cm^{-1} in this study is, in fact, comparable with those achieved via Raman spectroscopy in the fingerprint spectral region (accuracy of 90-100% in 800~1800 cm^{-1}) for *in vivo* detection of cervical dysplasia [100, 102]. Hence, HW Raman spectroscopy (2800~3700 cm^{-1}) together with PCA-LDA can be used to yield high diagnostic accuracy for *in vivo* detection of cervical precancer. However, this PCA-LDA model may be skewed to the detection of CIN 2 and 3 because CIN 2 and 3 subjects are dominant in the total dysplasia subjects. Hence, this model needs to be further validated on the dataset including comparable amount of CIN 1 subjects.

In addition, in the clinic diagnosis point of view, the discrimination between normal and dysplasia is not the whole colposcopic practice in clinic. The discrimination between LGSILs from HGSILs is also clinically desirable. The reason is that LGSILs have a high rate to regress to normal without any treatment [187] and however the current colposcopic diagnosis is likely to overestimate the LGSILs to be HGSILs. Meanwhile, discrimination between normal and LGSILs is even more challenging because some benign changes of cervix may also show some features of LGSILs [188]. For instance, HPV infection or inflammation may also exhibit vascular atypia which is the hallmark of higher grade lesion, and metaplasia may exhibit acetowhite due to the cellular changes [187, 188]. Therefore, we have also evaluated the performance of HW Raman spectroscopy in discriminating LGSILs from HGSILs, and LGSILs from

normal. For the former, 80.0% (8/10) of LGSILs and 97.2% (35/36) of HGSILs were distinguished correctly. For the latter, 80.0% (8/10) of LGSILs and 97.8% (45/46) of normal were distinguished correctly. These results show that HW Raman spectroscopy has the potential to differentiate LGSILs from both normal and HGSILs so as to eliminate the overassessment of colposcopic findings and protect LGSILs patients from unnecessary treatment. In addition, the results also encourage the use of HW Raman spectroscopy to resolve other challenges during colposcopic diagnosis, such as the low sensitivity to CIN 3, choice of biopsy site and inter-colposcopist variability [187, 188].

4.5 Conclusion

In conclusion, *in vivo* HW Raman spectra in the range of 2800~3700 cm^{-1} from normal and precancer cervical tissue can be acquired in 1 second using the rapid Raman system integrated with a ball lens fiber-optic Raman probe. Distinctive Raman spectral differences are observed for the first time between normal and dysplastic cervical tissue *in vivo* in the spectral range of 2800-3700 cm^{-1} which contains water, lipids and collagen signals. Good differentiation between normal and dysplastic cervical tissue can be achieved using HW Raman spectroscopy and PCA-LDA techniques. Due to the small data size, *in vivo* HW Raman measurements on a larger series of patients are currently in progress at NUH to further validate the clinical merits of HW Raman spectroscopy for the early diagnosis and detection of cervical precancer and cancer *in vivo* during clinical colposcopic examinations. In addition, clinical variables, such as subjects' age, race, menopausal status, chronic disease history, deserve to be taken into consideration to optimize the diagnostic model.

Chapter 5 *In Vivo* Diagnosis of Cervical Precancer Using NIR-excited Autofluorescence and Raman Spectroscopy

In this chapter, we investigated the feasibility of combining NIR autofluorescence (AF) and Raman spectroscopy for improving *in vivo* diagnosis and detection of cervical precancer. NIR Raman spectroscopy system coupled with a ball lens fiber-optic Raman probe was employed for *in vivo* Raman measurements on the cervix. A total of 92 *in vivo* Raman spectra (46 normal and 46 dysplasia) were acquired from the cervix of 46 patients under the guidance of colposcopy. Multivariate statistical techniques including principal components analysis (PCA) and linear discriminant analysis (LDA) were employed to develop diagnostic algorithms for differentiating precancer from normal tissue based on their spectral features. Classification results obtained from the leave-one patient-out, cross-validation of the PCA-LDA model based on the three spectral datasets (i.e., Raman, NIR AF, and the combined NIR AF and Raman) yielded diagnostic sensitivities of 91.3%, 93.5%, 93.5%, specificities of 95.7%, 87.0%, 95.7%, and accuracies of 93.5%, 90.2%, 94.6%, respectively, for precancer identification. Receiver operating characteristic (ROC) curves confirm that the most effective diagnostic algorithm can be derived from the combined NIR AF and Raman technique. This work demonstrates that the integrated NIR autofluorescence and Raman spectroscopy has a potential to provide an effective and accurate diagnostic scheme for non-invasive, *in vivo* diagnosis and detection of precancer in the cervix.

5.1 Introduction

Optical spectroscopic techniques, such as fluorescence spectroscopy, diffuse reflectance, infrared spectroscopy and Raman spectroscopy, have been comprehensively investigated for *ex vivo* and *in vivo* diagnosis of malignancies in various organs including the cervix [25, 46, 51, 72, 78, 98, 102, 138, 175, 189, 205, 249-251]. Light-induced autofluorescence (AF) spectroscopy has shown great promise in differentiating precancer from normal tissues presumably due to the fact that neoplastic transformation leads to the changes in tissue morphologic structures, optical properties (e.g., absorption and scattering coefficients), and the content and distribution of endogenous fluorophores in tissue [134, 138, 249, 250]. Raman spectroscopy is a molecular vibrational spectroscopic technique that is capable of probing biochemical and biomolecular changes of tissue associated with neoplastic transformation [51, 98, 189]. With the use of near-infrared (NIR) light as the excitation light source, NIR Raman spectroscopy holds significant advantages over other vibrational spectroscopy techniques (e.g., IR spectroscopy) in that water exhibits very low absorption at the working wavelength range, and tissues exhibit far less AF than that using visible light excitation, leading to an increased ratio of Raman signal to AF background [12]. Raman spectroscopy has revealed significant differences between normal and malignant tissues in various organs *ex vivo* and *in vivo* [25, 51, 78, 98, 100, 189], indicating the ability of Raman technique for cancer diagnosis. However, most studies have evaluated tissue Raman spectroscopy in isolation without considering the concomitant intense AF background from tissue. We recently found that NIR AF background still contained useful diagnostic information and retained the diagnostic ability, and hence combining Raman spectroscopy with NIR AF can enhance the diagnostic accuracy in skin tumor models [175]. In this study, we further extended the

NIR-excited tissue AF and Raman spectroscopy into a clinical setting to evaluate its utility for *in vivo* diagnosis of cervical precancer. A rapid-acquisition NIR Raman system coupled with a ball lens Raman probe developed in house [190] was used to acquire *in vivo* NIR Raman spectra associated with AF background from cervical tissue during clinical colposcopic examination. Multivariate statistical techniques, such as principal components analysis (PCA) and linear discriminant analysis (LDA), were employed to compare the diagnostic performance of Raman, NIR AF, and the combined Raman and NIR AF spectroscopy for classification of cervical precancer *in vivo*.

5.2 Materials and Methods

5.2.1 NIR Autofluorescence and Raman Instrumentation

The NIR AF and Raman instrumentation used in this study has been described in detail in Chapter 3.

5.2.2 Patients

The patient information has been given in the section **Patients** in Chapter 4.

5.2.3 Data Preprocessing

Under the 785-nm laser excitation, the raw Raman spectra acquired *in vivo* from cervix in the 800~1800 cm^{-1} range comprised prominent tissue NIR AF background, weak tissue Raman scattering signals, and noise. Hence, the raw Raman spectra were smoothed by a first-order, 5-pixel Savitzky Golay filter for noise reduction [178]. The broad NIR AF background was fit by a 5^{th} order modified polynomial [28, 51]. This polynomial was then subtracted from the raw Raman spectrum to yield the tissue Raman spectrum alone. Hence, three datasets including the raw Raman spectra

(combined Raman and NIR AF), NIR AF spectra (i.e., polynomials), and Raman spectra (i.e., background-subtracted raw Raman spectra) were generated, and utilized for multivariate statistical analysis.

5.2.4 Multivariate Statistical Analysis

PCA-LDA was employed for developing diagnostic algorithms and has been described in detail in the section **Multivariate Statistical Analysis** in Chapter 2.

5.3 Results

5.3.1 Spectral Feature Analysis

Figure 5.1 shows the mean *in vivo* spectra ± 1 standard deviation (SD) of (a) raw Raman spectra (i.e., combined NIR AF and Raman), (b) NIR AF spectra, and (c) Raman spectra acquired from normal (n=46) and precancer (n=46) tissues in the cervix. It was observed that both raw Raman spectra (Fig. 5.1(a)) and background NIR AF (Fig. 5.1(b)) exhibited significantly higher intensity in precancer tissue as compared to normal tissue (paired 2-sided Student's t-test, $p<0.00l$). On the contrary, Raman spectra showed a reduction in the overall intensity in dysplasia tissue as compared to normal tissue. However, due to the large inter- and/or intra-subjects variation and overlapping in spectral intensities illustrated by SD of all *in vivo* tissue spectra acquired (Fig. 5.1), tissue diagnosis simply based on the spectral intensity alone was precluded. It is also challenging in clinical settings to standardize the absolute intensities of spectral measurements. Therefore, in this study, we focused on analyzing the spectral variation in line-shape and relative intensity of NIR AF and Raman spectra associated with tissue neoplastic changes through normalizing each spectrum to the integrated area under the spectrum.

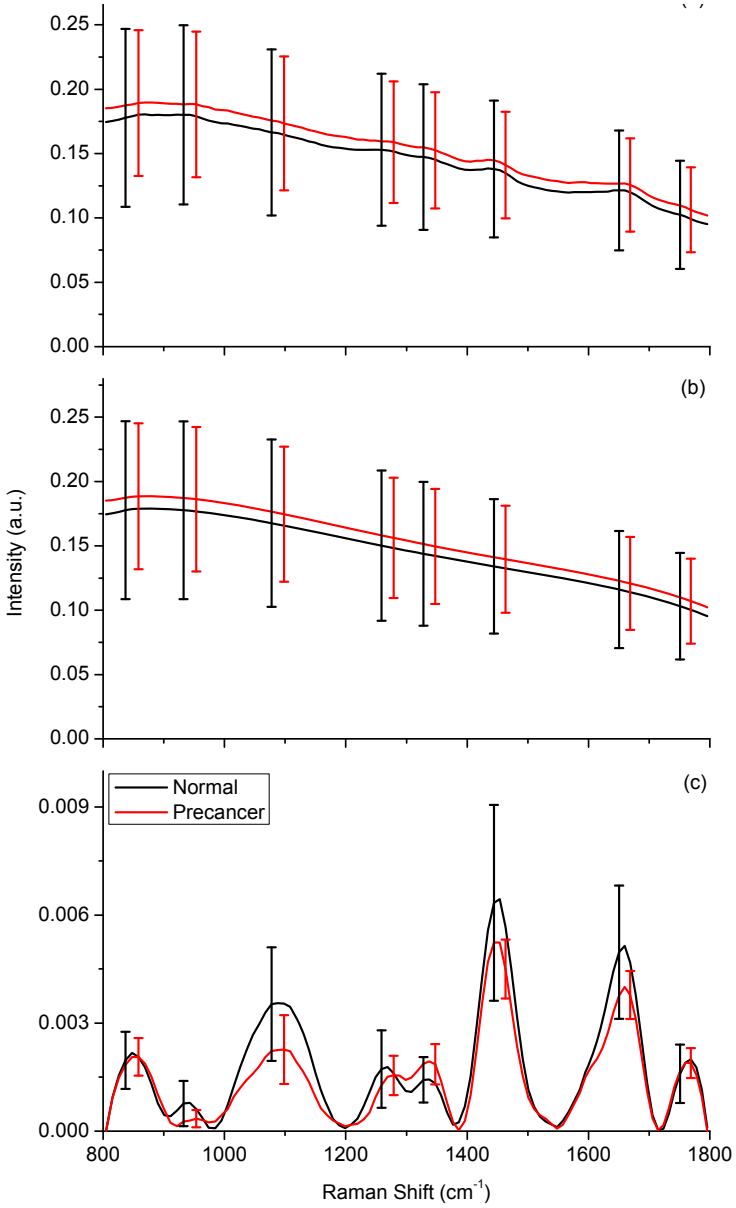


Figure 5.1 (a) Mean *in vivo* raw Raman spectra (combined Raman and NIR AF spectra) ± 1SD; (b) Mean *in vivo* NIR AF spectra (5^{th}-polynomials) ± 1SD; and (c) Mean *in vivo* Raman spectra (background-subtracted) ± 1SD from normal (n=46) and precancer (n=46) cervical tissue, respectively.

It was observed that raw Raman was a superimposition of weak Raman signal on relatively intense autofluorescence signal. Raman spectra of both normal and precancer tissues showed primary Raman bands in the vicinities of 850, 943, 1087, 1269, 1337, 1450, 1659 and 1765 cm^{-1} as shown in Fig. 5.1(c). The primary Raman peaks above can be tentatively assigned to different biochemical bonds as following: 850 to C-C

stretching of collagen, 943 to C-C stretching of protein, 1087 to phosphate ion stretching and C-N stretching of protein, 1269 to C-N stretching and N-H bending of amide III, 1337 to CH_2CH_3 wagging of proteins and nucleic acids, 1450 to CH_2 bending of proteins and lipids, 1659 to C=O stretching of Amide I and 1765 cm^{-1} to C=O stretching of phospholipids according to literature [51, 73, 102, 175, 189]. Therefore, Raman spectral variations in line shape and intensity associated with tissue premalignant transformation can be correlated to the changes in biochemical composition, accounting for the diagnostic potential of NIR Raman spectroscopy for cervical precancer. Although no distinctive shift in Raman peak position occurred between normal and precancer tissues, significant relative peak intensity variation can be observed in the intensity-normalized Raman spectra (data not shown).

On the other hand, NIR AF showed apparent spectral line-shape variation between normal and precancer. For example, NIR AF intensity declined from 800 to 1800 cm^{-1} more gently in dysplasia tissue than in normal tissue. More spectral variations were elucidated by the NIR AF spectra ratio of normal to precancer (data not shown), which declined steeply from 800 to 850 cm^{-1}, then gently from 850 to 1350 cm^{-1}, and eventually kept increasing until 1800 cm^{-1}. The spectral features above were subsequently utilized by PCA-LDA for developing diagnostic algorithms for differentiating dysplasia from normal tissues.

5.3.2 PCA-LDA and ROC Analysis

Figure 5.2 shows the first three diagnostically significant PC loadings yielded by PCA on the three datasets of: (a) raw Raman spectra, (b) NIR AF spectra and (c) Raman spectra. PC loadings revealed the correlation of each PC with original spectral data. For instance, PC 2 calculated from Raman spectra were negatively correlated with

Raman peak at 1087 cm^{-1} and positively correlated with Raman peaks at 1269 and 1450 cm^{-1} with loading value of 0.83, 0.78 and 0.6, respectively. This implies that PC 2 contains the spectral features of these three peaks. Moreover, PC loadings showed different patterns between NIR AF and Raman spectra. In particular, PC loadings of Raman spectra were characterized by narrow peaks and troughs while PC loadings generated from NIR AF spectra were comprised by a few broad bands. Meanwhile, some of the PC loading features above can also be observed in the PC loadings of raw Raman spectra. For example, PC 3 of raw Raman spectra showed an overall line-shape similar to that of NIR AF and similar peaks and troughs (i.e., trough at 1069 and peak at 1308 cm^{-1}) as that of Raman spectra. This suggests that PCA on raw Raman spectra can extract the diagnostic information originating from both NIR AF and Raman spectra. The total variance covered by the PCs above was 6.7%, 7.4%, and 12.3% for the combined NIR AF and Raman, NIR AF and Raman spectra, respectively, indicating the amount of the diagnostic information contained in the respective spectral datasets for differentiating between normal and precancer tissues.

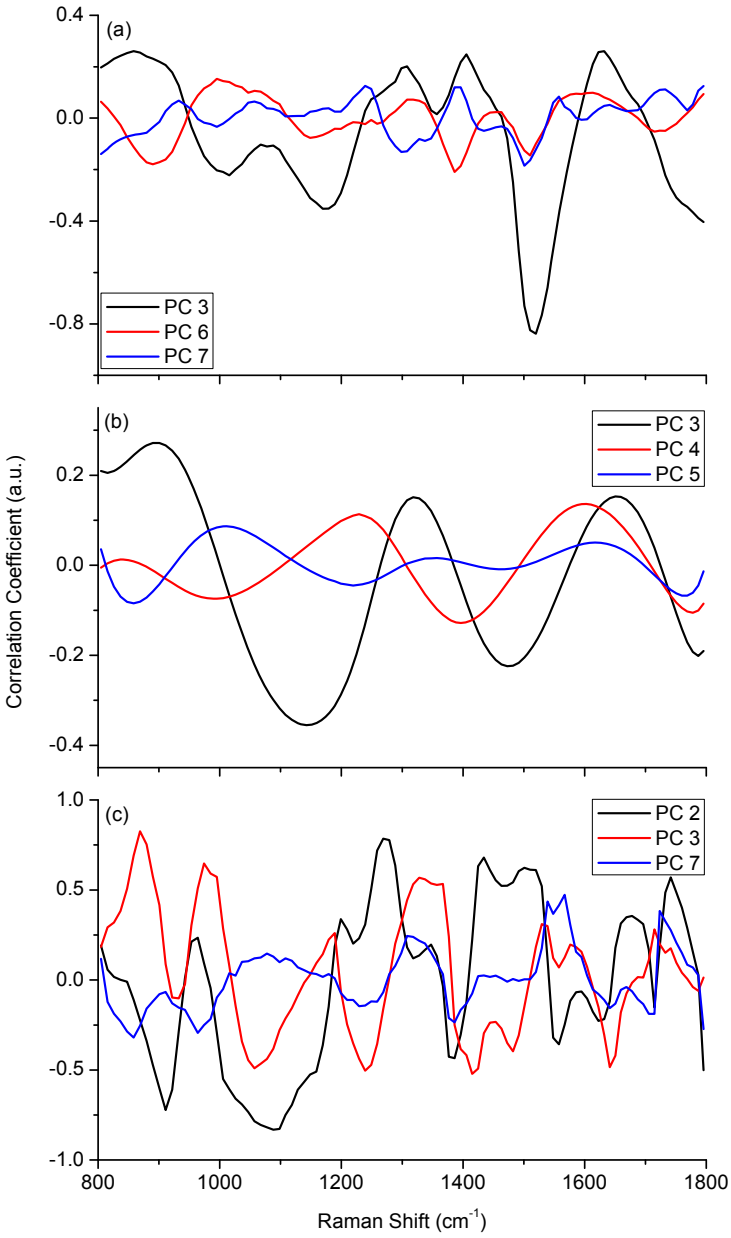


Figure 5.2 Examples of the diagnostically significant principal components (PCs) calculated from (a) raw Raman spectra (PC3, 4.8%; PC6, 1.5%; PC7, 0.4%;), (b) NIR autofluorescence spectra (PC3, 3.2%; PC4, 2.5%; PC5, 1.7%;) and (c) Raman spectra (PC2, 7.5%; PC3, 4.5%; PC7, 0.3%), respectively.

To illustrate the use of PC scores for the diagnostic classifications of the three spectroscopic techniques, Fig. 5.3 shows examples of scatter plots of the diagnostically significant PC scores for normal and precancer tissues: (a) raw Raman spectra (PC3 *vs* PC6), (b) NIR AF spectra (PC3 *vs* PC5), (c) Raman spectra (PC3 *vs* PC7). Simple straight lines can classify precancer from normal tissues with sensitivities of 84.8%,

80.4% and 87.0%; specificities of 84.8%, 73.9% and 78.3% for the combined NIR AF and Raman, NIR AF and Raman spectra, respectively. This indicates the diagnostic potential of the significant PC scores yielded by the three spectral datasets.

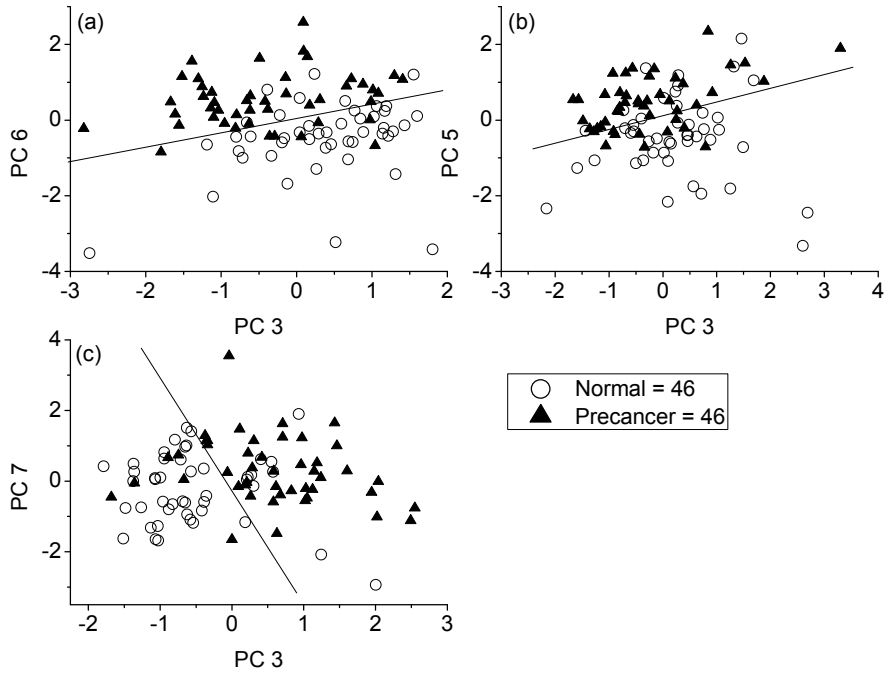


Figure 5.3 Correlations between the diagnostically significant PCs scores for normal and precancer cervical tissue classification: (a) raw Raman spectra, PC3 *vs* PC6, (b) NIR AF spectra, PC3 *vs* PC5, (c) Raman spectra, PC3 *vs* PC7. The separation lines (PC6 = 0.38×PC3+0.04; PC5 = 0.36×PC3+0.11; PC7 = −3.21×PC3−0.29) as diagnostic algorithms separate precancer from normal cervical tissue with sensitivities of 84.8% (39/46), 80.4% (37/46) and 87.0% (40/46); specificities of 84.8% (39/46), 73.9% (34/46) and 78.3% (36/46) using the three spectral datasets of raw Raman spectra (combined NIR AF and Raman spectra), NIR AF and Raman, respectively.

To incorporate all significant spectral features, LDA was employed to develop tissue diagnostic algorithms by utilizing all the statistically significant PC scores. Then, the classification accuracies achieved by the PCA-LDA diagnostic algorithms were considered as the criterion for evaluating the performances of different spectral datasets in tissue diagnosis. Figure 5.4 shows the posterior probabilities of each spectrum belonging to normal and precancer tissue groups as calculated for (a) combined NIR AF and Raman spectra, (b) NIR AF spectra, and (c) Raman spectra, respectively. The diagnostic sensitivities of 93.5% (43/46), 93.5% (43/46) and 91.3%

(42/46); specificities of 95.7% (44/46), 87.0% (40/46), and 95.7% (44/46); and accuracies of 94.6% (87/92), 90.2% (83/92) and 93.5% (86/92), respectively, can be achieved by using the combined NIR AF and Raman, NIR AF, and Raman techniques. It was found that the combined NIR AF and Raman achieved the highest sensitivity as NIR AF, and the highest specificity as Raman and consequently the optimal overall accuracy.

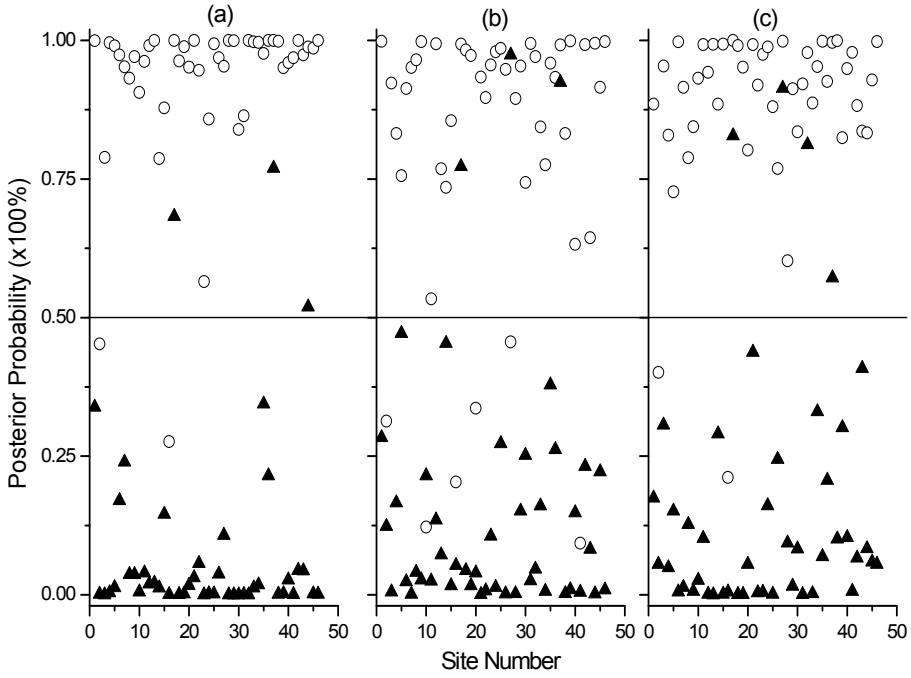


Figure 5.4 Scatter plots of the posterior probability of belonging to normal and precancer categories calculated from the datasets of (a) combined NIR AF and Raman, (b) NIR AF, and (c) Raman spectra, respectively, using the PCA-LDA-based spectral classification with the leave-one patient-out, cross-validation method. The corresponding sensitivity, specificity and accuracy are: (a) 93.5% (43/46), 95.7% (44/46), and 94.6% (87/92); (b) 93.5% (43/46), 87.0% (40/46), and 90.2% (83/92); (c) 91.3% (42/46), 95.7% (44/46), and 93.5% (86/92), respectively, using the combined NIR AF and Raman, NIR AF, and Raman techniques.

To further evaluate the performance of the PCA-LDA-based diagnostic algorithms on the three spectral datasets (i.e., combined NIR AF and Raman, NIR AF, and Raman), ROC curves were generated from the posterior probability distributions of different types of tissue (Fig. 5.4) as shown in Fig. 5.5. The areas under the ROC curves are 0.996, 0.945 and 0.972 for the combined NIR AF and Raman, NIR AF, and Raman

techniques, respectively. This further confirms the improvement in the diagnosis achieved by the combined NIR AF and Raman which contains both NIR AF and Raman spectral signatures as compared to either of NIR AF and Raman spectra alone.

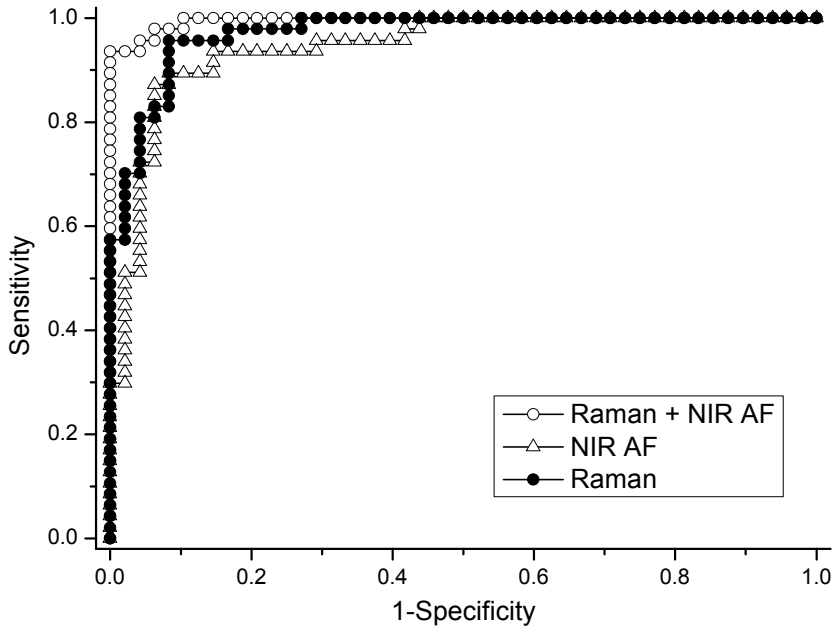


Figure 5.5 Receiver operating characteristic (ROC) curves of discrimination results for the combined NIR AF and Raman spectra, NIR AF, and Raman spectra, respectively. The integration areas under the ROC curves are 0.996, 0.945, and 0.972, respectively, for the combined NIR AF and Raman, NIR AF, and Raman techniques.

5.4 Discussion

Considerable research efforts have been made to develop advanced optical modalities for the noninvasive *in vivo* detection of cervical precancer during colposcopic examination. Optical spectroscopic techniques, such as fluorescence under the ultraviolet (UV) light or short-wavelength visible (VIS) light excitation, diffuse reflectance, light scattering and NIR Raman spectroscopy, have prevailed to date for cervical precancer and cancer diagnosis [102, 144, 148]. Among the different spectroscopic techniques, fluorescence, diffuse reflectance and light scattering are supposed to be complementary in detecting tissue neoplastic changes; therefore, they are combined for further improving the diagnosis of cervical precancer as compared to

the diagnosis by either of the three techniques alone [148]. Subsequently, Huang et al suggested that NIR AF and Raman also provide complementary information for cancer diagnosis in an *in vitro* study on mice skin model [175]. In this study, we attempted for the first time to combine NIR AF and Raman to improve and facilitate the *in vivo* diagnosis of cervical precancer. The diagnostic results produced by PCA-LDA diagnostic model calculated from the combined NIR AF and Raman, NIR AF and Raman demonstrate that although fairly good classification accuracy can be achieved by either of NIR AF and Raman alone, the diagnostic accuracy can be further improved by combining NIR AF with Raman.

To date, limited work has been conducted to explore NIR AF technique for tissue discrimination, probably due to the fact that most of the recognized fluorophores in tissue are not highly fluorescent under NIR excitation as NIR light is located in the tail part of the fluorophores' excitation spectrum [252]. Consequently, NIR AF spectrum is quite disparate from typical tissue fluorescence spectrum at the excitation of UV/VIS light. This also holds true for our study, showing that tissue NIR AF spectrum exhibited broad band almost like straight line instead of a composite of narrow spectral bands. Our study also showed that tissue NIR AF was more intense in precancer tissue as compared to normal tissue. This may be attributed to the elevated porphyrins in precancer tissue. The reason is probably that porphyrins exhibits an intense emission with significant absorption of NIR light [252]. Moreover, endogenous porphyrins is elevated in malignant or neoplastic tissues, due to its affinity with cancer cells [253] and hypervascularization which is more common in malignant tissues [254]. On the other hand, Raman spectra of cervix in our study showed similar prominent Raman peaks as reported in literature [102]. Except Raman band at 1337 cm^{-1}, all the prominent Raman bands showed a decreased intensity in precancer tissue than in

normal tissue as reported in our previous study [175]. The intensity increase of Raman band at 1337 cm^{-1} may result from the increased nuclear content in precancerous tissue [51]. Besides the absolute intensity change, distinctive difference in relative band intensity was also found between normal and precancer tissue (data not shown). For instance, Raman bands at 1087 and 1659 cm^{-1} due to collagen exhibited lower intensity in precancer tissue than in normal tissue while Raman bands at 1450 and 1765 cm^{-1} due to phospholipids showed otherwise [51, 98]. The relative peak intensity change can be related to biochemical changes (i.e., reduced collagen content and increased phospholipids content) associated with tissue neoplastic transformation [72, 98, 254], forming the base of NIR Raman spectroscopic diagnosis of cervical precancer. The diagnostic accuracy yielded by PCA-LDA model (i.e., 90.2% for NIR AF and 93.5% for Raman) further confirms the diagnostic ability of both NIR AF and Raman for cervical dysplasia.

However, in most of the past and current NIR Raman studies for tissue diagnosis, the NIR AF background is usually discarded as non-diagnostic information. Our results show that Raman spectrum was superior to NIR AF background in terms of specificity (95.7% for Raman *vs* 87.0% for NIR AF). This may be explained by the fact that Raman spectrum shows a bigger amount of fingerprint spectral features representing molecular structure and composition of tissues [51, 255]. However, the diagnostic sensitivity (91.3%) by Raman technique was lower than that of NIR AF (93.5%). The relatively higher sensitivity of NIR AF to tissue dysplastic changes may be owed to the intense fluorescence emission from tissue, particularly compared to weak Raman scattering. Hence, one diagnostic technique with a sensitivity and specificity as high as that of fluorescence and Raman, respectively, is desirable for improving *in vivo* diagnosis of cervical dysplasia. The diagnostic results based on PCA-LDA model show

that the combined NIR AF and Raman achieved the highest sensitivity as NIR AF alone and the highest specificity as Raman alone, respectively. This indicates that combining NIR AF and Raman is an efficient choice of method to utilize the diagnostic features of both NIR AF and Raman for developing diagnostic algorithm so as to further improve or optimize the tissue diagnosis. The combination-induced improvement may be further explained as the following. It is known that tissue is composite of various fluorophores (e.g. keratin, collagen, lipids, etc.), of which each has a different autofluorescence spectral pattern under the excitation of specific wavelength. Meanwhile, most of the biomolecules (e.g. proteins, lipids, DNA, etc.) in tissue are Raman-active, showing fingerprint Raman spectral features. Consequently, the raw Raman spectral data (composite Raman and AF signatures) allows the examination of the endogenous fluorophores and morphological structures in tissue (autofluorescence) and the biomolecular structure and composition of tissue (Raman). Thereby, it can provide more diagnostic information for tissue diagnosis and characterization, resulting in an improvement in tissue diagnosis. Despite the small case number, the performances of the diagnostic algorithms derived from the combined NIR AF and Raman, NIR AF, and Raman spectra by using PCA-LDA with leave-one-out cross validation have formed a foundation and shown the potential for the *in vivo* diagnosis of cervical precancer.

In addition, to better understand the spectroscopic diagnosis, we also evaluated the interaction of the NIR excitation light with cervical tissue by using Monte Carlo (MC) method. The propagation of 785-nm light in cervical tissue and the light interaction with cervical tissue were simulated based on a two-layer (i.e., epithelium and stroma) cervical tissue optical model [229]. The results show that more than 90% of collected NIR AF and Raman photons originated in the epithelium. This suggests that our

Raman probe coupled with a ball lens has a good depth-selectivity to the epithelium of cervical tissue and as a result can enhance the detection of NIR AF and Raman from the epithelium. Note that precancer changes are usually confined in the epithelium. Hence, the use of our ball lens Raman probe may improve the sensitivity for detecting cervical dysplasia by NIR AF and Raman.

5.5 Conclusion

In this study, we demonstrated that combining NIR AF and Raman can improve the diagnosis of cervical precancer as compared to either of NIR AF and Raman techniques alone. This implies that NIR AF and Raman are not necessarily competing approaches but could be complementary by providing informations about different aspects of cervical tissue under investigation. In addition to spectroscopy combination, these two techniques may be more expected to be combined in another form of NIR AF image-guided Raman spectroscopy in some clinic procedures. For example, in surgical resection of malignant tissue, it is desirable to delineate the tumor margin accurately. In this case, NIR fluorescence imaging can serve to rough the malignant tissue quickly and subsequently Raman spectroscopy can further confirm the tumor margin. Therefore, it can be concluded that Raman spectroscopy combined with NIR AF spectroscopy or imaging shows promise as an effective and clinically useful diagnostic schema for the detection of precancer and cancer as well as an aid for the cancer therapy. Further investigation on a larger series of patients in which the diagnostic algorithms are tested prospectively on new cases is ongoing to confirm these results. Meanwhile, the development of an *in vivo* NIR autofluorescence imaging system is also in progress for future imaging-guided Raman spectroscopic diagnosis.

Chapter 6 Combining NIR Raman, UV/VIS Autofluorescence and Diffuse Reflectance Spectroscopy for Improving Cervical Precancer Detection

Optical spectroscopic techniques, such as near-infrared (NIR) Raman, ultraviolet/visible (UV/VIS) autofluorescence and reflectance spectroscopies, have shown a great potential for the detection of precancer and cancer in the cervix. However, these techniques are mostly employed separately and very little work has been done to combine these three different techniques for optimizing the diagnosis of pre-neoplastic (i.e., benign) and neoplastic changes (i.e., LGSILs and HGSILs) in cervical tissues. Therefore, in this study, we aimed at combining these three optical spectroscopic modalities to further improve the discrimination of cervical dysplasia. Eighty Raman (785-nm excitation), fluorescence (405-nm excitation) and reflectance spectra were acquired from the same batch of cervical biopsied tissues, including 24 spectra from 9 benign, 34 spectra from 12 LGSILs and 22 spectra from 9 HGSILs tissues. Trinary classification results yielded by principal components analysis (PCA) and linear discriminant analysis (LDA) together with the leave-one-out, cross-validation method showed that Raman spectroscopy can predict benign, LGSILs and HGSILs cervical tissues correctly with accuracies of 83.3%, 94.1% and 100%, respectively, and the corresponding total accuracy is 92.5% which is higher than than that of fluorescence (81.3%) and reflectance (87.5%). Furthermore, the prediction was improved through combining NIR Raman with UV/VIS fluorescence and reflectance, showing accuracies of benign-100%, LGSILS-97.1% and HGSILs-100%, overall-98.8%. This demonstrated that NIR Raman, UV/VIS fluorescence and reflectance can provide complementary diagnostic information and hence be integrated

together for optimizing the optical spectroscopic diagnosis of cervical dysplasia as compared to the individual use of either of Raman, fluorescence and reflectance.

6.1 Introduction

Optical spectroscopic techniques, such as Raman, fluorescence and reflectance spectroscopy, have emerged as promising methods for *ex vivo* and *in vivo* diagnosis of malignancies in various organs [51, 78, 98, 100, 119, 131, 134, 138, 150, 151, 175, 189, 190, 250, 251, 256]. In particular for cervix, Anita and coworkers succeeded in measuring Raman spectra in fingerprint region (1000~1800 cm^{-1}) with 785-nm excitation light from cervix *ex vivo* and *in vivo* and achieved a fairly good differentiation among normal, and dysplastic tissues [98, 100-102]. Recently, we demonstrated the feasibility of using Raman spectra in HW region for discriminating cervical dysplasia *in vivo* with sensitivity of 93.5% and specificity of 97.8% [190]. Meanwhile, Rebecca and coworkers have done an extensive work on applying fluorescence spectroscopy for the detection of cervical precancer, such as optimizing excitation light wavelength as well as diagnostic algorithms [119, 127, 131, 134, 257]. Reflectance spectroscopy has also been developed for the *in vivo* diagnosis of cervical precancer [150, 151]. All the literature work above demonstrated that Raman, fluorescence and reflectance spectroscopies have the potential for the detection of cervical precancer with advantages, such as higher accuracy, noninvasive, fast and subjective as compared to conventional method (i.e., pap smear and colposcopy).

Although these three different techniques can do similar diagnostic work, the biomarkers which are utilized by them for diagnosis are different. Specifically, Raman spectroscopy is a vibrational spectroscopic technique and can provide molecular fingerprint information about biological tissues. Therefore, it is capable of probing the

changes in tissue biochemical composition associated with malignant transformation [51, 98, 189]. The major biomarkers sensible for Raman spectroscopy include proteins, lipids, nucleic acids and carbohydrates [51, 98, 189]. In comparison, fluorescence spectroscopy is based on the transition of electronic state of atoms and molecules. It can mainly detect the changes in a limited number of fluorophores (i.e., tryptophan, flavin adenine dinucleotide (FAD), reduced form of Nicotinamide adenine dinucleotide (NADH), collagen and porphyrins) in tissues [258]. Reflectance spectrum is a measure of tissue scattering and absorption properties, which are associated with the morphology and architecture of tissues. From the introduction above, it can be concluded that Raman, fluorescence and reflectance make diagnosis through providing insights into different aspects of tissues. The diagnostic information (i.e., biochemical composition and morphology) characteristic of biological tissues could be complementary and thus used in combination for optimizing the optical spectroscopic diagnosis of pre-cancer and cancer. However, to date, most of optical spectroscopic diagnosis work was performed with an individual use of either of these three different techniques rather than in combination [259]. Therefore, in this study, we aimed to combine near-infrared (NIR) Raman, ultraviolet/visible (UV/VIS) fluorescence and reflectance spectroscopy together for the first time to improve the early diagnosis of cervical dysplasia. In addition, intrinsic fluorescence was retrieved for diagnostic algorithm development as measured fluorescence is inevitably distorted by tissue rescattering and reabsorption [260]. Multivariate statistical techniques, including principle component analysis (PCA) and linear discriminant analysis (LDA), were employed to develop diagnostic algorithms. The combination was implemented based on the strategies which produce the diagnostic decision with the use of the diagnostic results yielded by either of Raman, fluorescence and reflectance alone.

6.2 Materials and Methods

6.2.1 Spectroscopy Instrumentation

Figure 6.1 depicts the trimodal spectroscopic measurement setup. It comprises Raman, fluorescence and reflectance spectroscopy. Raman spectroscopy system has been described in the section **Raman Instrumentation** in Chapter 2. It mainly consists of a spectrum stabilized 785-nm diode laser (B&W TEK, Newark, DE), a transmissive imaging spectrograph, a NIR-optimized, back-illuminated, and deep-depletion charge-coupled device (CCD) detector (Holospec f/2.2, Kaiser Optical Systems, Trento, NJ), and a specially developed bifurcated fiber-optic Raman probe.

Fluorescence spectroscopy mainly comprises a spectrometer (QE65000, Ocean Optics, Dunedin, FL), a 405-nm diode laser (LDCU12/7318, 50 mW, Power Technology, AR, USA) as the excitation light source and a bifurcate fiber-optic probe (QR400-7-VIS-NIR, working wavelength region: 400~2500 nm, Ocean Optics, Dunedin, FL). The probe is constructed with a single 400-μm fiber for excitation and six surrounding 400-μm fibers for collection. The excitation laser light is filtered by 405-nm band pass filter (LD01-405/10-25, bandwidth=10 nm, Semrock, NY, USA) and then coupled into the central fiber of the probe via a SMA (SubMiniature version A) connector. The emitted fluorescence photons are collected by the surrounding collection fibers and then filtered by 405-nm long pass filter (BLP01-405R-25, cut off at 418 nm, Semrock, NY, USA) to block the back-scattered excitation light prior to being fed into spectrometer.

Diffuse reflectance spectroscopy is built by removing the band pass and long pass filters from fluorescence spectroscopy setup and replacing the laser with a tungsten halogen light source (LS-1, Ocean Optics, Dunedin, FL). The illumination light is directly coupled into the central fiber of the probe via a SMA connector. The back-scattered light photons are collected by the surrounding six collection fibers. The collected photons are eventually fed into the spectrometer to create the reflectance spectrum.

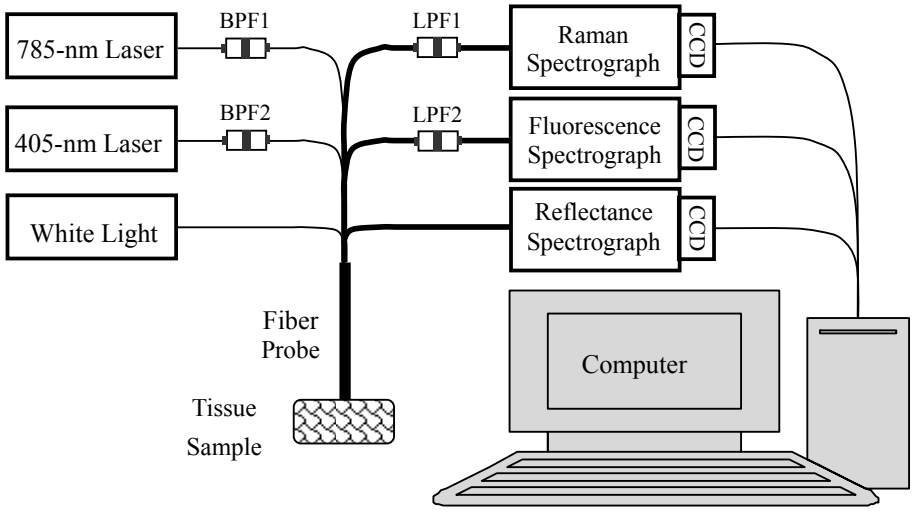


Figure 6.1 Schematic of the trimodal spectroscopy system. BPF1: 785-nm band pass filter; BPF2: 405-nm band pass filter; LPF1: 800-nm long pass filter; LPF2: 405-nm long pass filter;

6.2.2 Cervical Tissue Samples

Tissue samples measured has been described in the section **Cervical Tissue Samples** in Chapter 2.

6.2.3 Spectroscopic Measurement

All the tissue spectroscopic measurements were done in a light-tight box which can eliminate most of the interference from the environmental lights (e.g., computer

monitor, light indicators). The tissue samples were mounted onto aluminum foil for the measurement. Raman, fluorescence and reflectance spectra were acquired from the same tissue sample in tandem. The acquisition times were 1 to 3 seconds for Raman and fluorescence measurement and 0.1 to 0.2 seconds for reflectance measurement depending on the signal to noise ratio. All Raman and fluorescence spectra were corrected for the wavelength-dependent intensity response function of the system using a standard lamp (RS-10, EG&G Gamma Scientific, San Diego, CA). Diffuse reflectance measurement includes the acquisition of reference spectrum from Diffuse Reflectance Standard (WS-1, Ocean Optics, Dunedin, FL), dark spectrum with the light off and tissue spectrum. The reflectance spectrum is created by the calculation based on the following equation:

$$REL = (S - D)/(R - D) \tag{6.1}$$

where S is the light intensity from sample, D is the dark spectrum, and R is the reference spectrum.

6.2.4 Data Preprocessing

1) Raman: Raman spectral data preprocessing has been described in the section **Data Preprocessing** in Chapter 2.

2) Fluorescence/Reflectance: Fluorescence spectra in the region of 430~700 nm and reflectance spectra in the region of 400~700 nm were truncated for analysis. The spectrum was noise-reduced and normalized in the same way as Raman spectrum.

3) Retrieve intrinsic fluorescence: A simple model was used to retrieve the intrinsic

tissue fluorescence signal from the measured fluorescence signal by correcting the effect of tissue reabsorption and rescattering. This model is to divide the measured fluorescence spectrum by the reflectance spectrum acquired from the same tissue to a power of k as illustrated by the following equation [260]:

$$FL_i = FL_m \Big/ REL^k \tag{6.2}$$

where FL_i is the retrieved intrinsic fluorescence, FL_m is the measured fluorescence, and REL is the corresponding diffuse reflectance spectrum. The power k is the ratio between the path lengths of fluorescence and diffuse reflectance light which is the function of light wavelength due to the wavelength-dependent scattering properties of tissues. In this model, it is assumed to be independent of light wavelength to simplify the model. The power k is optimized by minimizing the hemoglobin-absorption effect on the fluorescence spectrum. It is noted that fluorescence and diffuse reflectance spectra appear to have two dips which are centered at the similar wavelength positions (i.e., 542 and 576 nm) for fluorescence and diffuse reflectance. With the assumption that the dip area represents the extent of the hemoglobin absorption, the power k is optimized by minimizing the dip area. The dip area is approximated by the triangle covering the dip. Two triangles are determined by two sets of three points (520, 542, 560 nm) and (560, 576, 590 nm) for the two dips at 542 and 576 nm, respectively. Three different powers are generated by minimizing 542 nm dip area, 576 nm dip area and the sum of the areas of dips at both 542 and 576 nm. The corresponding fluorescence is denoted as intrinsic fluorescence 1, 2 and 3, respectively.

6.2.5 Multivariate Statistical Analysis

The multivariate statistical methods have been described in the section **Multivariate Statistical Analysis** in Chapter 2.

6.2.6 Strategies of Combining Raman, Fluorescence and Reflectance

Four different strategies were employed to combine Raman, fluorescence and reflectance spectroscopy to explore the possibility of further improving the diagnosis compared to the diagnosis by either of Raman, fluorescence and reflectance alone. The first one is to make the final diagnosis with the majority of the predictions by different techniques, so called majority-determined prediction. The other three combination strategies were implemented by utilizing the posterior probabilities of each case belonging to benign, LGSILs and HGSILs, which were produced by these three techniques. These three strategies are to determine the combined posterior probability of each case belonging to benign, LGSILs and HGSILs with the product, maximum and mean of its posterior probabilities generated by Raman, fluorescence and reflectance diagnosis, which are also called product-determined, max-determined and mean-determined predictions, respectively. The final diagnosis was determined by the highest combined posterior probability of belonging to benign, LGSILs and HGSILs. Note that one of fluorescence datasets (i.e., measured and intrinsic fluorescence) with the highest accuracy, one of reflectance datasets (i.e., non-normalized and normalized reflectance) with higher accuracy and Raman are selected for the combination.

6.3 Results

6.3.1 NIR Raman

The results of NIR Raman have been presented in the section **Results** in Chapter 2.

6.3.2 UV/VIS Fluorescence

Figure 6.2 depicts the mean normalized fluorescence spectra of benign, LGSILs and HGSILs: (a) non-corrected fluorescence, (b) intrinsic fluorescence 1, (c) intrinsic fluorescence 2 and (d) intrinsic fluorescence 3. Non-corrected fluorescence spectra showed two obvious absorption dips, which however almost disappeared in intrinsic fluorescence 1, 2 and 3. Apart from this, non-corrected fluorescence spectra and intrinsic fluorescence spectra looked very alike. Obvious intensity difference among benign, LGSILs and HGSILs was observed, especially in the region of 475~530 nm, while relatively small difference in the regions of 430~475 nm and 600~700 nm. Moreover, the three different power k optimization criteria exhibited similar correction extent illustrated by almost the same spectral shape. The corresponding k values are 0.518 ± 0.217, 0.603 ± 0.198 and 0.548 ± 0.213 for intrinsic fluorescence 1, 2 and 3, respectively.

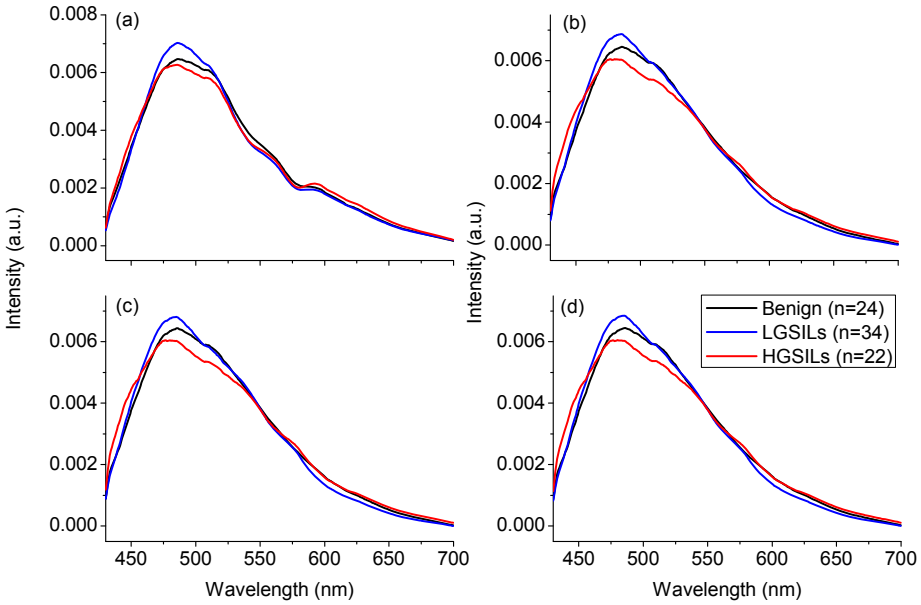


Figure 6.2 The mean normalized fluorescence spectra of benign (n=24), LGSILs (n=34) and HGSILs (n=22): (a) non-corrected fluorescence, (b) intrinsic fluorescence 1, (c) intrinsic fluorescence 2, and (d) intrinsic fluorescence 3.

To test whether correcting the effect of tissue absorption and scattering can improve the tissue diagnosis and also find out the optimal correction criterion, PCA-LDA was employed to develop the diagnostic algorithms to evaluate the efficacy of the various correction methods. Figure 6.3 shows the correlation of the first five diagnostically significant ($p<0.1$) PCs with the wavelength variables: (a) PC1, PC3, PC4, PC5 and PC12 for non-corrected fluorescence, accounting for 69.5%, 5.68%, 1.05%, 0.8% and 0.01% of the total variance, respectively; (b) PC1, PC2, PC5, PC11 and PC12 for intrinsic fluorescence 1, accounting for 71.3%, 14.1%, 1.05%, 0.02% and 0.01% of the total variance, respectively; (c) PC1, PC2, PC5, PC11 and PC12 for intrinsic fluorescence 2, accounting for 69.1%, 14.1%, 1.04%, 0.02% and 0.01% of the total variance, respectively; (d) PC1, PC2, PC5, PC11 and PC12 for intrinsic fluorescence 3, accounting for 70.7%, 14.1%, 1.04%, 0.02% and 0.01% of the total variance, respectively. It was observed that PC1 covers most (~70%) of the total variance for the four different fluorescence datasets. It also showed a big similarity in the shape and coefficient value among different datasets. However, there still existed a slight

difference on PC1 between non-corrected fluorescence and intrinsic fluorescence, which was that a small peak at the vicinity of 576 nm appeared in PC1 of non-corrected fluorescence while disappearing in PC1 of intrinsic fluorescence. This peak may be correlated to the blood absorption dip at 576 nm and consequently the difference may be explained by the correction of the blood absorption.

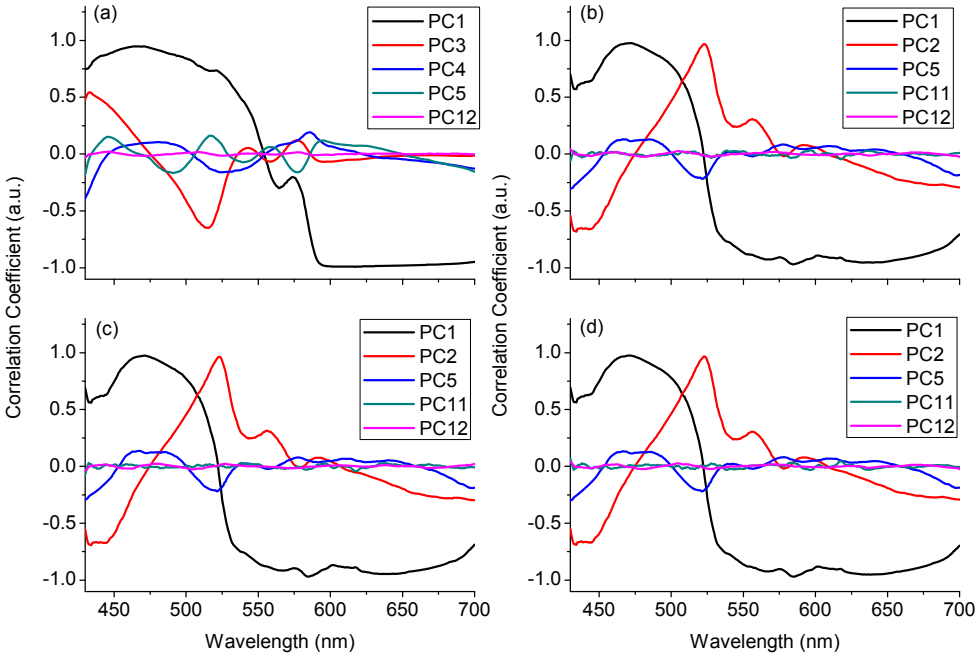


Figure 6.3 Examples of the first five diagnostically significant principal components (PCs) with p-value<0.1: (a) PC1, PC3, PC4, PC5 and PC12 for non-corrected fluorescence; (b) PC1, PC2, PC5, PC11 and PC12 for intrinsic fluorescence 1; (c) PC1, PC2, PC5, PC11 and PC12 for intrinsic fluorescence 2; (d) PC1, PC2, PC5, PC11 and PC12 for intrinsic fluorescence 3.

To develop diagnostic algorithms, LDA was performed with all the significant PCs as input. Two LD functions were produced for the three-group classification. Figure 6.4 shows scatter plots of two LD function weights for the four different spectral datasets. It was observed that the data points belonging to the same group mostly assembled together, indicating the efficacy of PCA-LDA for characterizing benign, LGSILs and HGSILs. In addition, a better separation of HGSILs cluster from the clusters of benign and LGSILs occurred compared to the separation between benign and LGSILs. Figure 6.5 shows the posterior probabilities of each case belonging to benign, LGSILs and

HGSILs produced by the PCA-LDA diagnostic models. The distribution of the HGSILs was better separated from the distribution of benign and LGSILs, especially for the intrinsic fluorescence datasets.

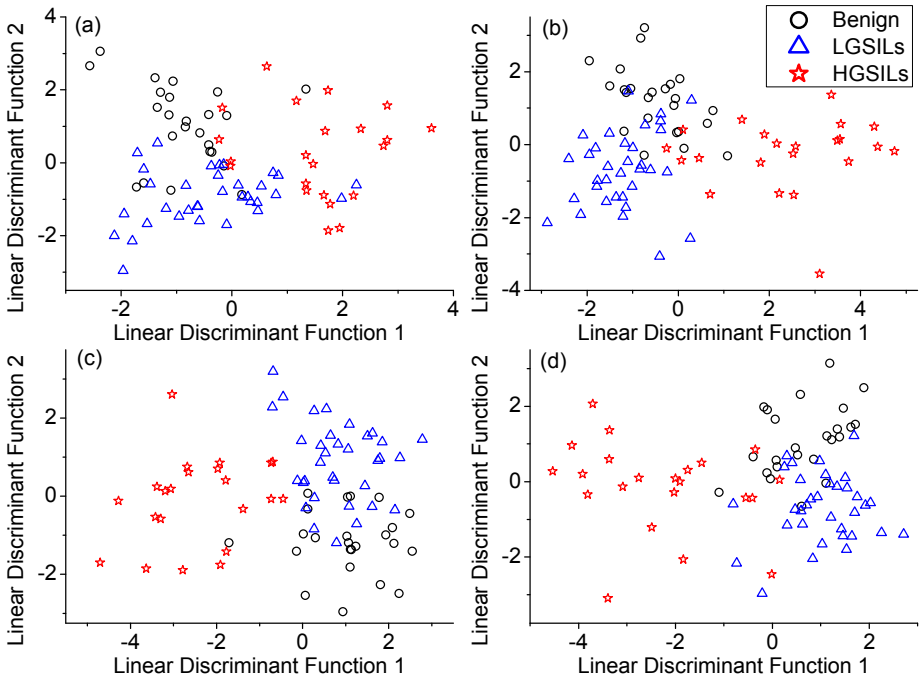


Figure 6.4 Scatter plots of two LD function weights for benign (n=24), LGSILs (n=34) and HGSILs (n=22) tissues tested with leave-one-out cross-validation: (a) non-corrected fluorescence; (b) intrinsic fluorescence 1; (c) intrinsic fluorescence 2; (d) intrinsic fluorescence 3. Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

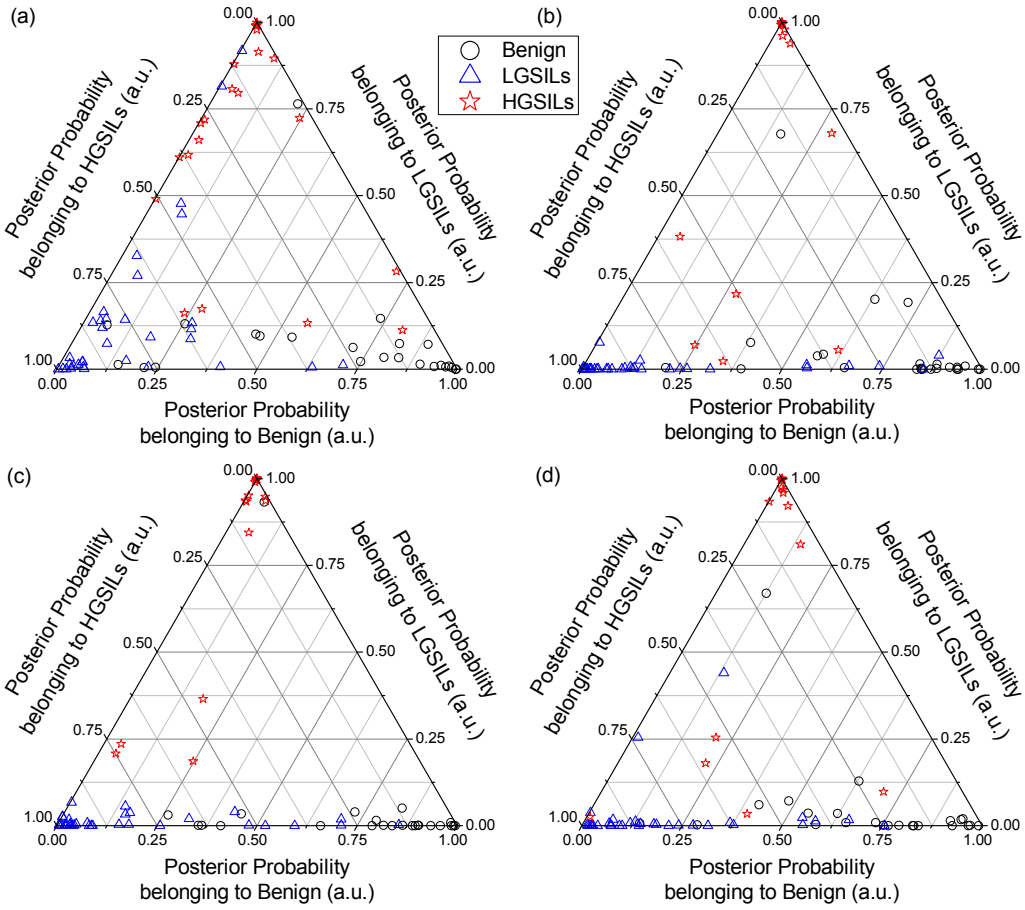


Figure 6.5 Two-dimensional ternary plots of the posterior probabilities of each spectrum belonging to benign (n=24), LGSILs (n=34) and HGSILs (n=22), respectively, using the PCA-LDA-based spectral classification with leave-one-out, cross-validation method based on the four different dataset: (a) non-corrected fluorescence; (b) intrinsic fluorescence 1; (c) intrinsic fluorescence 2; (d) intrinsic fluorescence 3. Each vertex of the triangle represents a 100% confidence that the tissue is benign, LGSILs or HGSILs. Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

The predications are summarized in table 6.1. The sensitivities, specificities and accuracies calculated from the statistics data in table 6.1 are described in table 6.2. An increase was found in the diagnostic sensitivity and total accuracy of benign and HGSILs by intrinsic fluorescence compared to non-corrected fluorescence. In contrast, intrinsic fluorescence did not show any significant improvement in the identification of LGSILs. The overall accuracy for benign, LGSILs and HGSILs was 77.5% (62/80) for non-corrected fluorescence spectra while an increased rate of 81.3% (65/80) is achieved for intrinsic fluorescence 1, 2 and 3.

Table 6.1 Classification results of fluorescence-prediction for the three cervical tissue groups yielded by the PCA-LDA diagnostic algorithms tested with leave-one-out cross-validation method.

	Tissue Type	Fluorescence-predicted Tissue Type			Total
			Dysplasia		
		Benign	LGSILs	HGSILs	
Non-correction	Benign	17	6	1	24
	LGSILs	2	29	3	34
	HGSILs	3	3	16	22
Intrinsic fluorescence 1	Benign	20	3	1	24
	LGSILs	6	28	0	34
	HGSILs	1	4	17	22
Intrinsic fluorescence 2	Benign	18	5	1	24
	LGSILs	5	29	0	34
	HGSILs	0	4	18	22
Intrinsic fluorescence 3	Benign	20	3	1	24
	LGSILs	5	28	1	34
	HGSILs	1	4	17	22

Table 6.2 The sensitivity, specificity and accuracy of fluorescence-prediction for the three cervical tissue groups yielded by the PCA-LDA diagnostic algorithms tested with leave-one-out cross-validation method.

		Fluorescence-predicted Tissue Type		
			Dysplasia	
		Benign	LGSILs	HGSILs
Non-correction	Sensitivity	70.8% (17/24)	85.3% (29/34)	72.7% (16/22)
	Specificity	91.1% (51/56)	80.4% (37/46)	93.1% (54/58)
	Accuracy	85.5% (68/80)	82.5% (66/80)	87.5% (70/80))
Intrinsic fluorescence 1	Sensitivity	83.3% (20/24)	82.4% (28/34)	77.2% (17/22)
	Specificity	87.5% (49/56)	84.8% (39/46)	98.3% (57/58)
	Accuracy	86.3% (69/80)	83.8% (67/80)	92.5% (74/80)
Intrinsic fluorescence 2	Sensitivity	75.0% (18/24)	85.3% (29/34)	81.8% (18/22)
	Specificity	91.1% (51/56)	80.4% (37/46)	98.3% (57/58)
	Accuracy	86.3% (69/80)	82.5% (66/80)	93.8% (75/80)
Intrinsic fluorescence 3	Sensitivity	83.3% (20/24)	82.4% (28/34)	77.3% (17/22)
	Specificity	89.3% (50/56)	84.8% (39/46)	96.6% (56/58)
	Accuracy	87.5% (70/80)	83.8% (67/80)	91.3% (73/80)

Due to the fact that three intrinsic fluorescence datasets showed the same overall accuracy, it was hard to judge which correction can provide the optimal diagnostic performance. Also, it is desirable to further confirm the improvement in the diagnosis by correcting the tissue absorption and scattering effect. Therefore, three-way ROC

analysis was performed on the posterior probabilities calculated from all the four datasets. Figure 6.6 depicts the three-way ROC surfaces for the four different spectral datasets. The ROC surfaces for intrinsic fluorescence spectra (Fig. 6.6(b) to (d)) exhibited an obvious increase in the region of TPR of HGSILs (0~1.0) and TPR of benign (0.75~1). The volume under the ROC surface (VUS) was calculated, which was 0.675, 0.754, 772 and 0.74 for non-corrected fluorescence, intrinsic fluorescence 1, 2 and 3, indicating that intrinsic fluorescence 2 was the optimal one.

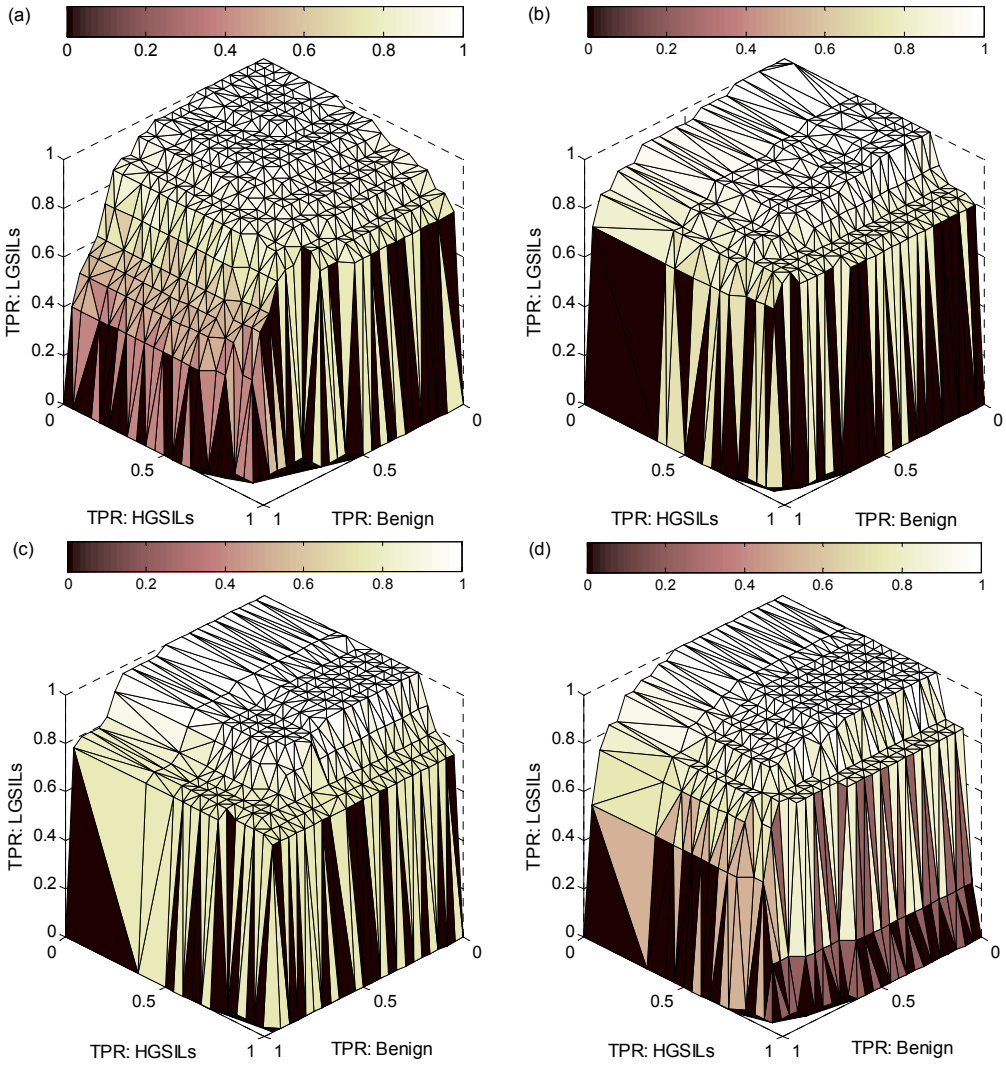


Figure 6.6 Three-dimensional view of the ROC surface calculated from the posterior probabilities belonging to benign, LGSILs and HGSILs based on the four different datasets: (a) non-corrected fluorescence; (b) intrinsic fluorescence 1; (c) intrinsic fluorescence 2; (d) intrinsic fluorescence 3. The corresponding volumes under ROC surface are: 0.675, 0.754, 0.772 and 0.74, indicating the rank of the diagnostic performance based on different datasets.

6.3.3 Diffuse Reflectance

Figure 6.7 shows the mean non-normalized (Fig. 6.7(a)) and normalized (Fig. 6.7(b)) spectra of benign, LGSILs and HGSILs. Overall, benign tissues appeared to exhibit the highest reflectance, HGSILs showed the weakest reflectance and LGISLs showed the reflectance in between. The reflectance difference between benign and dysplasia (LGSILs and HGSILs) was larger than that between LGSILs and HGSILs, especially in the blood absorption dip area in the central part of the reflectance spectrum. To avoid the misdiagnosis caused by the inter- and/or intra-patient variation, all the spectra were normalized to the integrated area under the spectrum as shown in Fig. 6.7(b). It was observed that the spectra were altered continuously in the overall slope from benign to LGSILs to HGSILs. For instance, the reflectance spectrum of HGSILs seemed the steepest while the gentlest slope for benign and LGSILs in between. These spectral differences among different types of tissues indicate the potential of reflectance spectroscopy for the characterization of benign, LGSILs and HGSILs cervical tissues.

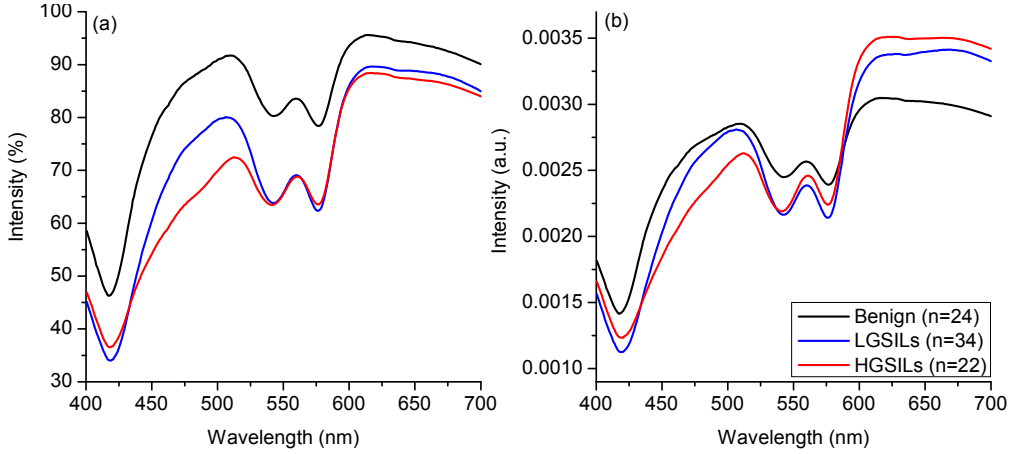


Figure 6.7 Mean reflectance spectra of benign (n=24), LGSILs (n=34) and HGSILs (n=22): (a) non-normalized spectra; (b) normalized spectra. Black: benign; Blue: LGSILs; Red: HGSILs.

To utilize those spectral differences for tissue diagnosis, PCA was firstly used to reduce the data dimension from 80×383 to 80×79. Then, one-way analysis of variance

(ANOVA) was performed on the PC scores to find out those diagnostically significant PCs with p-value<0.1. Figure 6.8 shows the correlation of the first four significant PCs with the wavelength variable: (a) PC1, PC4, PC5, PC6 for non-normalized spectral dataset, accounting for 74.3%, 1.33%, 0.33% and 0.05% of the total variance, respectively; (b) PC1, PC3, PC4 and PC5 for normalized spectral dataset, accounting for 78.2%, 8.98%, 0.76% and 0.18% of the total variance, respectively. Obviously, PC1 covered the majority of the total variance for both non-normalized and normalized spectral datasets, and however showed a distinctive difference in the correlation with wavelength variables for the two datasets. This may be due to the intensity feature which was gotten rid of by normalization. Also, the correlation curves appeared to contain some features, of which a rough counterpart could be found in the reflectance spectrum, such as two absorption dips in the region of 520~600 nm. This indicates that the spectral features remain in the PCs after dimension reduction.

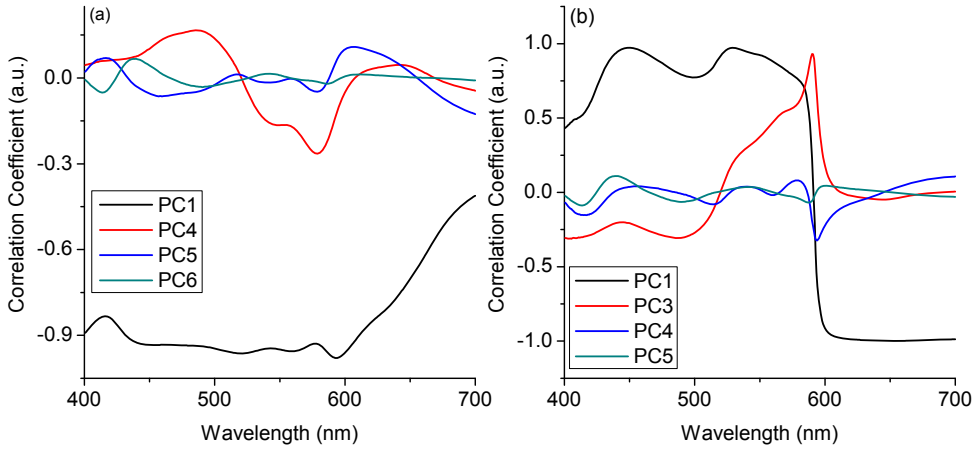


Figure 6.8 Examples of the first four diagnostically significant principal components (PCs) with p-value<0.1: (a) PC1, PC4, PC5 and PC6 for non-normalized spectral dataset; (b) PC1, PC3 PC4 and PC5 for normalized spectral dataset.

To develop diagnostic algorithms, LDA was carried out with the significant PC scores as input. Two linear discriminant (LD) functions were generated for tri-group classification. Figure 6.9 depicts the scatter plots of the LD functions for (a)

non-normalized spectra and (b) normalized spectra. It was noticed that the spectra from the same type of tissue was mostly assembled into the same cluster for non-normalized spectra and clustering was improved for normalized spectra. Figure 6.10 presents the posterior probabilities of each spectrum belonging to benign, LGSILs and HGSILs.

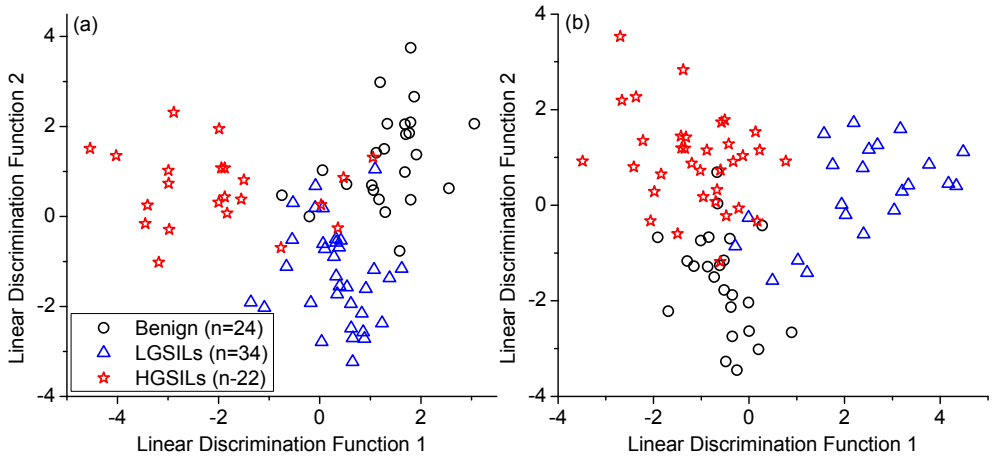


Figure 6.9 Scatter plots of two linear discrimination function weights for benign (n=24), LGSILs (n=34) and HGSILs (n=22) tissues tested with leave-one-out cross-validation: (a) non-normalized spectral dataset; (b) non-normalized spectral dataset; Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

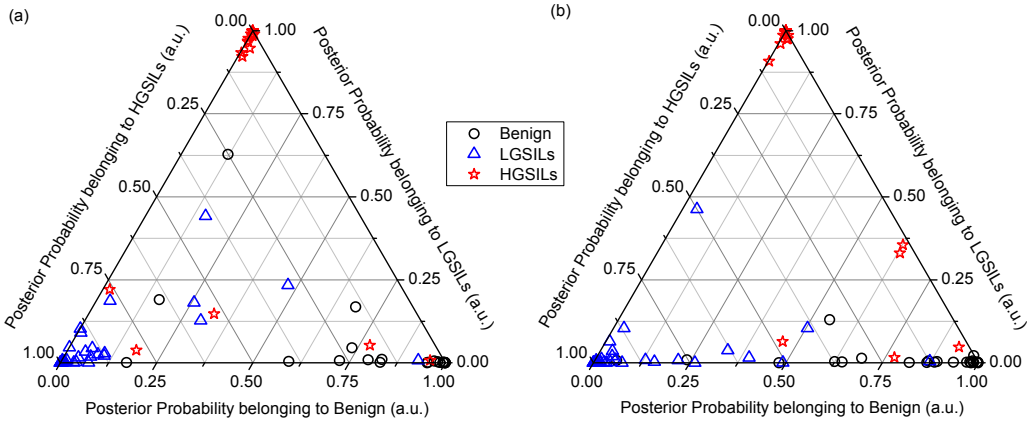


Figure 6.10 Two-dimensional ternary plots of the posterior probabilities belonging to benign, LGSILs and HGSILs, respectively, using the PCA-LDA-based spectral classification with leave-one-out, cross-validation method based on the dataset: (a) non-normalized spectra; (b) normalized spectra. Each vertex of the triangle represents a 100% confidence that the tissue is benign, LGSILs or HGSILs. Key: (○ in black) benign; (Δ in blue) LGSILs; (☆in red) HGSILs.

The classification results yielded by PCA-LDA model are summarized in table 6.3 and the corresponding sensitivities and specificities are tabulated in table 6.4. Twenty-one

benigns, thirty-one LGSILs and seventeen HGSILs were predicted correctly by non-normalized spectra. Normalized spectra yielded the same correct predication rate for benign and HGSILs and improved the prediction of LGSILs by one case. Although both non-normalized- and normalized-based LDA models showed the same accuracy in predicting benign and HGSILs, the misclassification was different. For example, three benigns were misclassified, including two as LGSILs and one as HGSILs for non-normalized spectra. In contrast, three benigns were all misclassified as LGSILs for normalized spectra. Hence, Fig. 6.11 elucidates the prediction variation between these two datasets. The LDA models based on non-normalized and normalized spectra got into an agreement in the prediction of 82.5% of the tissue spectra.

Table 6.3 Reflectance-prediction for the three cervical tissue groups yielded by the PCA-LDA diagnostic algorithms tested with leave-one-out cross-validation method.

	Tissue Type	Reflectance-predicted Tissue Type			Total
			Dysplasia		
		Benign	LGSILs	HGSILs	
Non-normalized	Benign	21	2	1	24
	LGSILs	2	31	1	34
	HGSILs	2	3	17	22
Normalized	Benign	21	3	0	24
	LGSILs	2	32	0	34
	HGSILs	4	1	17	22

Table 6.4 The sensitivity, specificity and accuracy of reflectance-prediction for the three cervical tissue groups yielded by the PCA-LDA diagnostic algorithms tested with leave-one-out cross-validation method.

		Reflectance-predicted Tissue Type		
			Dysplasia	
		Benign	LGSILs	HGSILs
Non-normalized	Sensitivity	87.5% (21/24)	91.2% (31/34)	77.3% (17/22)
	Specificity	92.9% (52/56)	89.1% (41/46)	96.6% (56/58)
	Accuracy	91.3% (73/80)	90.0% (72/80)	91.3% (73/80))
Normalized	Sensitivity	87.5% (21/24)	94.1% (32/34)	77.3% (17/22)
	Specificity	89.3% (50/56)	91.3% (42/46)	100.0% (58/58)
	Accuracy	88.8% (71/80)	92.5% (74/80)	93.8% (75/80)

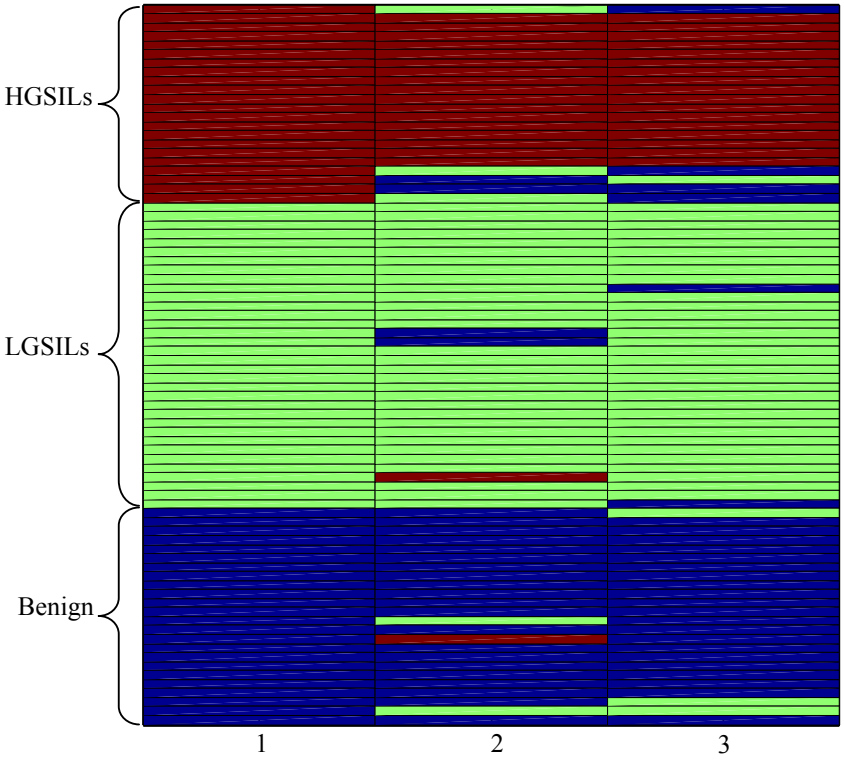


Figure 6.11 Color coded prediction results by the PCA-LDA model based on non-normalized and normalized spectral datasets. column 1: histology classification; column 2: non-normalized reflectance; column 3: normalized reflectance. Each grid represents one case. Blue: benign; Green: LGSILs; Brown: HGSILs.

To further confirm the improvement in the diagnosis achieved by using the normalized spectral data, three-way receiver operating characteristic (ROC) testing was performed. Figure 6.12 shows the three-way ROC surface. The volumes under the ROC surface were 0.773 and 0.78 for non-normalized and normalized spectra data, respectively, indicating that normalized spectra achieve a better diagnosis performance.

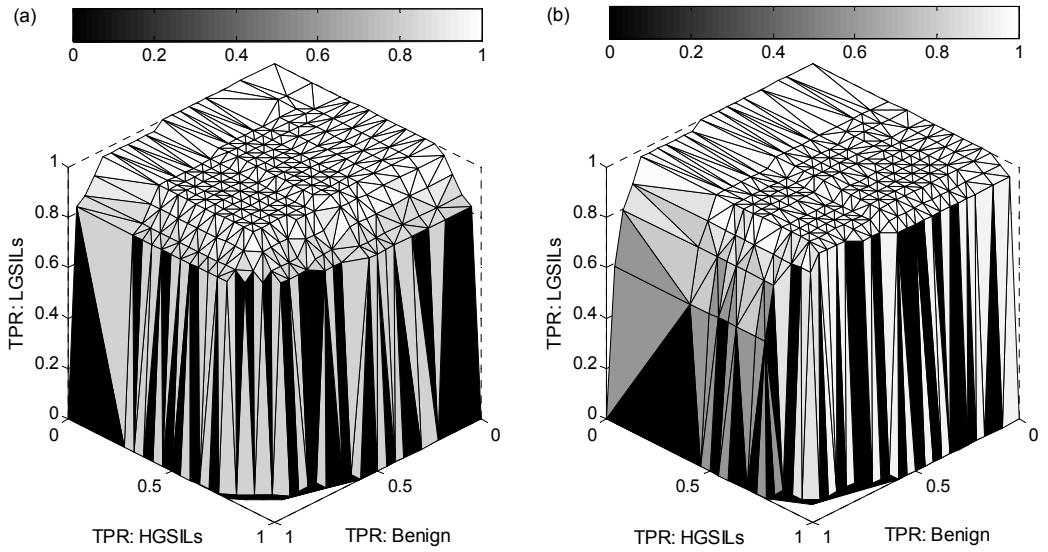


Figure 6.12 Three-dimensional view of the ROC surface calculated from the posterior probabilities belonging to benign, LGSILs and HGSILs based on the datasets: (a) non-normalized reflectance spectra; (b) normalized reflectance spectra. The corresponding volumes under ROC surface are: 0.773 and 0.78, indicating the rank of the diagnostic performance based on the two different datasets.

6.3.4 Compare and Combine NIR Raman, Fluorescence and Reflectance

The results above indicated that intrinsic fluorescence 2 is the best of the four different fluorescence spectra datasets with VUS of 0.772, and normalized reflectance is better than non-normalized reflectance with VUS of 0.78. A comparison of the VUS of Raman, intrinsic fluorescence 2 and normalized reflectance produced a rank of the diagnostic performance of these three different techniques, which is in an order of Raman, reflectance and fluorescence from best to worst.

To investigate the complementarity among Raman, fluorescence and reflectance, normalized reflectance and intrinsic fluorescence 2 were included in the combination with Raman as they showed better diagnostic performance than the remaining reflectance and fluorescence datasets. Figure 6.13 maps the prediction results for individual spectroscopic methods and combined methods. It was found that Raman, fluorescence and reflectance got in an agreement on predicting 70% of spectra, implying a potential improvement room for the diagnosis of cervical dysplasia. The

majority-determined was eventually excluded due to the unique prediction on four cases by three different datasets. The diagnostic accuracies achieved by product-determined, max-determined, mean-determined were 97.5% (78/80), 98.8% (79/80) and 97.5% (78/80), respectively, which were higher than either of Raman, fluorescence and reflectance alone. This demonstrates that combining the different spectroscopic techniques have the potential to further improve the diagnosis, of which max-determined prediction is the optimal one.

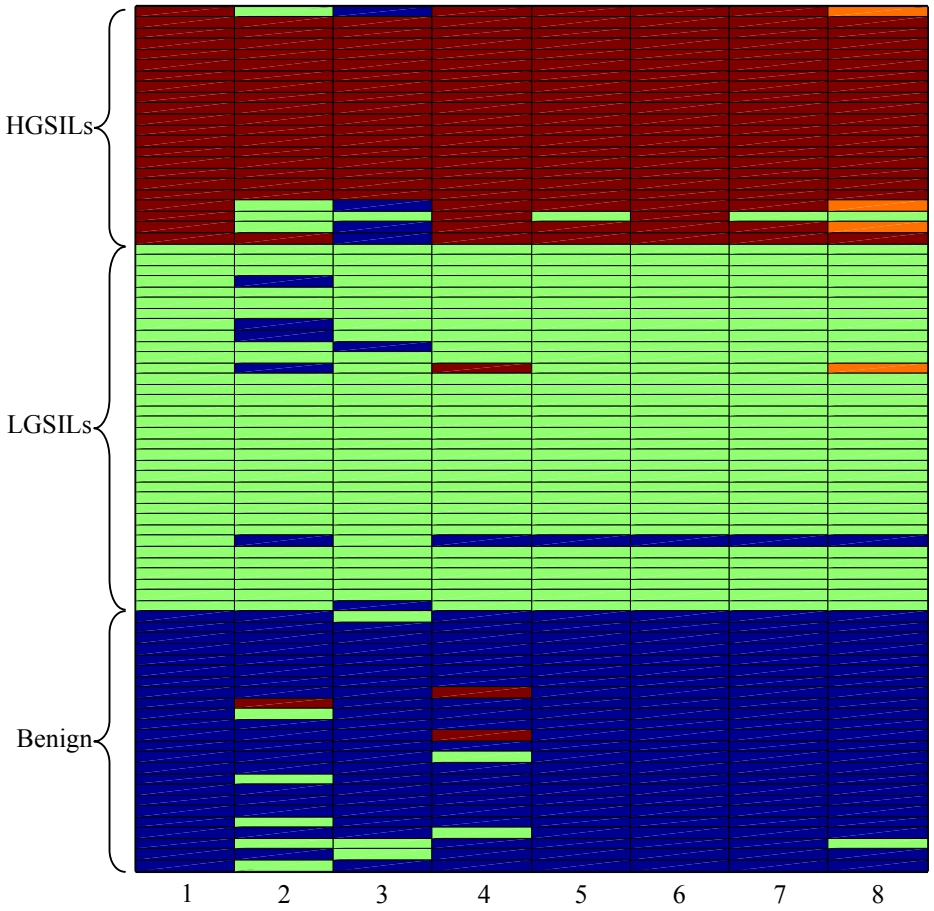


Figure 6.13 Color coded prediction results by combining intrinsic fluorescence 2, normalized reflectance and Raman. column 1: histology prediction; column 2: fluorescence prediction; column 3: reflectance prediction; column 4: Raman prediction; column 5: product-determined prediction; column 6: max-determined prediction; column 7: mean-determined prediction; column 8: majority-determined prediction; Blue: benign; Green: LGSILs; Brown: HGSILs; Orange: disagreement on predicting this case by fluorescence, reflectance and Raman.

6.4 Discussion

In this study, we evaluated the diagnostic capacities of three common spectroscopic techniques, including NIR Raman, UV/VIS fluorescence and reflectance spectroscopy, for cervical precancer and moreover explored the feasibility of further improving the diagnosis by combining these three techniques. Thus, NIR Raman, UV/VIS fluorescence and reflectance spectra were measured from the same set of cervical biopsied tissues, including three different types of pathologies (i.e., benign, LGSILs and HGSILs). This allowed us to make a direct pairwise comparison of diagnostic ability among Raman, fluorescence and reflectance as well as combine these three different techniques to optimize the diagnosis of cervical precancer. Good quality Raman, fluorescence and reflectance spectra were successfully acquired from cervical tissues. Significant spectral differences of Raman, fluorescence and reflectance were observed among benign, LGSILs and HGSILs. The diagnostic Raman spectral features have been discussed in detail in the section **Discussions** in Chapter 2. We discussed fluorescence and reflectance in detail in this chapter.

Fluorescence spectra were mainly characterized by a single broad band centered at 486 nm and two weak absorption dips at 542 and 576 nm as shown in Fig. 6.2(a). Note that fluorescence photon undergoes reabsorption and rescattering by those non-fluorescent absorbers and scatters in tissue. This can be illustrated in part by the two weak absorption dips. As a result, the measured tissue fluorescence spectrum is not the intrinsic fluorescence but the fluorescence distorted by tissue reabsorption and rescattering. This distortion may deteriorate the accuracy in predicting cervical dysplasia by fluorescence [261]. Besides, intrinsic fluorescence also facilitates the quantitative analysis of the concentration change of the fluorophores with tissue neoplastic transformation [138]. Therefore, it is desirable to investigate the diagnostic

ability exclusively of intrinsic fluorescence. In recent years, some methods have been proposed to retrieve intrinsic fluorescence from measured fluorescence by several groups [260, 262]. In this study, we employed a simple model to extract intrinsic fluorescence proposed by Diana et al [260]. The overall line shape of intrinsic fluorescence spectra remains similar to the measured fluorescence spectra, except the vanishing of the absorption dips, indicating the efficacy of this method on retrieving the intrinsic fluorescence from the measured fluorescence spectra.

The tissue fluorescence spectra at 405-nm excitation can be mostly attributed to NADH, FAD, collagen and porphyrins, which have been proven to be fluorescent and diagnostically relevant to cancerous changes in various human organs [117, 134, 138, 261, 263-265]. In particular for cervix, NADH and FAD dominate the epithelium fluorescence while collagen is the major contributor to stroma fluorescence [134, 137]. NADH is found to be increased while FAD is decreased in precancer tissues, which may be due to the increased metabolic activity associated with tissue premalignancy [117]. Collagen shows a decreased contribution to the tissue fluorescence in precancer tissues [117]. This may be explained by the increased collagenases which degrade the fluorescing collagen-crosslink in precancer tissues. This may also be partly due to the changes of tissue optical properties (e.g., thickening of epithelium) and morphologies (e.g., higher cellular density resulting from the increased nucleus to cytoplasm ratios) associated with dysplastic transformation. These changes above significantly attenuate the excitation light penetration and also obscure the tissue fluorescence photons emitted from the underlying stroma in precancer tissue as compared to normal cervical tissue [145, 192]. Our results show that HGSILs exhibited highly reduced fluorescence intensity. This may be due to the reduced FAD and collagen contribution. In comparison, fluorescence intensity above 600 nm was elevated in HGSILs as

compared to benign and LGSILs. This is probably attributed to the elevated porphyrins showing affinity with cancer cells, which is consistent with the reported finding in literature [115]. The classifications by PCA-LDA diagnostic models show that three groups of intrinsic fluorescence spectra yielded the same overall accuracy of 81.3% (65/80), which was higher than that achieved by measured fluorescence spectra (77.5% (62/80)). This indicates that intrinsic fluorescence has the potential to discriminate cervical dysplasia from benign tissues. The three-way ROC-VUS confirms its diagnostic potential and also suggests extracting intrinsic fluorescence with criterion 2 is optimal for tissue discrimination.

Reflectance spectra of three tissue groups (Fig. 6.7) were dominated by three prominent valleys, which were located in the vicinities of 418, 542 and 576 nm, respectively. These three valleys are recognized to result from the hemoglobin absorption and have been found in diffuse reflectance spectrum of a variety of human tissues [150, 151, 176, 260, 261, 266]. Moreover, the overall spectral intensity decreased continuously as the cervix tissue progresses from benign to HGSILs as shown in Fig. 6.7(a). This is consistent with the findings on cervix in the literature [150, 151]. This decrease in the reflectance of dysplasia can be accounted for by the increased hemoglobin absorption due to the fact that cancer tissue shows an increased angiogenesis with an increased microvessel density compared to normal or benign tissues [193]. In addition, the normalized reflectance spectrum (Fig. 6.7(b)) showed that the reflectance intensity increased in the region of 625~675 nm while decreased in the region of 450~500 nm with the progression of tissue from benign to HGSILs, indicating an increase in the spectrum slope. This slope change has been reported to be diagnostically valuable [151]. The reflectance spectral features above build a basis for reflectance spectroscopy coupled with multivariate statistical techniques (i.e.,

PCA-LDA) as a diagnostic method for cervical precancer. The classification results show that PCA-LDA diagnostic model calculated from normalized data achieved almost the same overall accuracy (70/80) except one LGSILs case improvement compared to non-normalized data (69/80). However, the misclassification distribution was different, leading to various specificities in the diagnosis of each type of tissues. For instance, compared to non-normalized data, normalized data showed an increased sensitivity and specificity for LGSILs while the same sensitivity and a lower specificity for benign. Three-way ROC-VUS revealed the improvement extent by normalized data (0.78 *vs* 0.773) as compared to non-normalized data.

Comparing the diagnostic accuracy and ROC-VUS calculated from Raman, fluorescence and reflectance, it was obvious that Raman outperformed fluorescence and reflectance in differentiating among benign, LGSILs and HGSILs while fluorescence showed the weakest diagnostic ability. The same diagnostic ranking of Raman, fluorescence and reflectance was reported on breast cancer [259]. The superior performance of Raman over fluorescence and reflectance may be explained as below. Raman spectroscopy is a molecular fingerprint technique and Raman spectrum of biological tissue contains many narrow Raman peaks which are assigned to different biomolecules. This allows Raman spectroscopy to probe the tissue biolomecular changes associated with tissue malignant transformation with high accuracy. In contrast, tissue fluorescence spectrum appears to be relatively broad, flat and featureless due to the overlap of the emission spectra from different fluorophores [175, 259]. This may account for the less diagnostic information and consequently the inferior diagnostic performance compared to Raman. As for reflectance spectroscopy, it provides insight into tissue absorption (i.e., haemoglobin) and scattering properties (i.e., reduced scattering coefficient) related to tissue morphological structures [150,

151, 176]. Obviously, the diagnostic information contained in reflectance spectrum is very limited, especially compared to Raman spectrum, leading to the relatively weaker diagnostic ability. As for the comparison between fluorescence and reflectance, discrepancy is found in literature [144, 149, 259, 260].

Furthermore, to our knowledge, we attempted to combine Raman, fluorescence and reflectance to improve the diagnosis of cervical dysplasia for the first time. Before this study, only fluorescence, diffuse reflectance and elastic scattering spectroscopy were included in the combination for improving the detection of cancer and precancer, [148, 176, 259, 260]. Our previous study demonstrated that NIR Raman and autofluorescence are complementary rather than competitive in cancer diagnosis [175]. Therefore, we got Raman involved in the combination with fluorescence and reflectance for improving the diagnosis of cervical precancer. The combination of different spectroscopy methods can be achieved mainly in two different ways, which are combination of spectra for developing diagnostic algorithm and combination of classifier (i.e., use the majority of classifiers by different methods), respectively. The first one was reported to succeed in combining fluorescence and reflectance for improving breast tissue discrimination [259]. Meanwhile, the improvement by the second one is also reported in literature. For example, Georgakoudi et al succeeded in improving the diagnosis of esophageal dysplasia by combing fluorescence, reflectance and light-scattering spectroscopy with this method [176]. More recently, Veld et al attempted to use the maximum, product and mean of each case's posterior probabilities yielded by fluorescence and reflectance to improve the diagnosis [260]. We used the same method to combine Raman, fluorescence and reflectance in our study. Our results demonstrated that combining Raman, fluorescence and reflectance has the potential to further improve the diagnosis of cervical dysplasia and max-determined prediction is

found to be the optimal combination strategy for improving the diagnosis of cervical dysplasia. The improvement resulting from the combination above can be accounted for by the fact that Raman, fluorescence and reflectance describe tissue in terms of different aspects of tissue as discussed above: Raman spectra provide the composition information of various biochemicals; fluorescence spectra probe a limited number of fluorophores and reflectance spectra provide the information about tissue morphology and architecture.

6.5 Conclusion

In this study, we evaluated the performance of NIR Raman, UV/VIS fluorescence and reflectance spectroscopy in the diagnosis of cervical dysplasia. The results showed that NIR Raman exhibited better diagnostic ability than UV/VIS fluorescence and reflectance, showing an overall diagnostic accuracy of 92.5% versus 81.3% (fluorescence) and 87.5% (reflectance). Then, we demonstrated that the diagnosis can be further improved by combining NIR Raman with UV/VIS fluorescence and reflectance with a significantly improved overall diagnostic accuracy of 98.8%. This implied that Raman, fluorescence and reflectance can provide complementary information for tissue diagnosis by providing different aspects of tissue pre-malignancy and malignancy. In addition, this combination can also create a more comprehensive understanding of the biochemical and morphological changes associated with tissue cancerous transformation, which may be helpful for treatment design.

Chapter 7 Conclusions and Future Work

7.1 Conclusions

The entire thesis is centered on NIR Raman spectroscopic diagnosis of the dysplasia in the cervix. The results and findings obtained from this thesis work demonstrate that NIR Raman spectroscopy has the potential to emerge as a promising technique for screening and diagnosing cervical dysplasia, and preventing cervical cancer.

First of all, we explored the potential of NIR Raman spectroscopy for the detection of cervical precancer *ex vivo*. An in-house developed NIR Raman spectroscopy was employed and a novel data acquisition program was developed on LabVIEW platform for Raman measurement on biopsied cervical tissues. Significant spectral differences were observed among various tissue pathologies (i.e., benign, LGSILs and HGSILs). For instance, a significant decrease occurred on the intensity of Raman peaks at 849 (C-C stretching of proline ring and ring breathing mode of tyrosine), 1125 (C-C stretching of lipid, C-N stretching of protein) and 1658 cm^{-1} (Amide I, C=O stretching mode of collagen and elastin) while an increased intensity for Raman peaks at 1004 (phenylanaline), 1339 (nucleic acid) and 1449 cm^{-1} (phospholipids). Empirical analysis (i.e., intensity ratio of Raman peaks) and multivariate statistical techniques (i.e., PCA-LDA) were used to develop diagnostic algorithms. The PCA-LDA diagnostic model can make correct prediction for 92.5% of the tissues (i.e., benign, LGSILs and HGSILs). The volume under three-ways ROC surface was 0.815, indicating the total diagnostic accuracy. In addition, tissue spectral model revealed the major origins of biochemical changes with tissue precancerous transformation, such as the reduced contribution from collagen and glycogen and elevated contribution from lipids (i.e., cholesterol and oleic acid) and DNA. All the results above demonstrate that NIR

Raman spectroscopy has the potential for cervical precancer.

The *ex vivo* study above encouraged an extension of our work to an *in vivo* study. A rapid portable NIR Raman spectroscopy system with a fiber-optic Raman probe was developed for the *in vivo* Raman measurement on the cervix. The fiber-optic Raman probe was coupled with a ball lens to enhance the detection of Raman signal originating from epithelial tissue, especially the epithelium. Monte Carlo method was employed to evaluate the performance of the ball lens Raman probe design. The Monte Carlo simulation results show that tissue Raman collection efficiency can be improved by properly selecting the refractive index and the diameter of the ball lens for the Raman probe design and the depth-selectivity of Raman measurements can also be improved by either increasing the refractive index or reducing the diameter of the ball lens. The simulation and experimental validation results demonstrate that this ball lens Raman probe design can facilitate the depth-selected Raman measurements of epithelial tissue *in vivo*.

With the use of the *in vivo* Raman system we developed, we explored the feasibility of NIR Raman spectroscopy in the HW region (2800~3700 cm^{-1}) for *in vivo* detection of cervical precancer as HW Raman is free of fiber optics' interference and is more intense than Raman scattering in fingerprint region (800~1800 cm^{-1}). Significant differences in Raman intensities of prominent Raman bands peaking at 2850 and 2885 cm^{-1} (CH_2 stretching of lipids), 2940 cm^{-1} (CH_3 stretching of proteins), and the broad Raman band of water (peaking at 3400 cm^{-1} in the 3100-3700 cm^{-1} range) were observed in normal and dysplasia cervical tissue. A PCA-LDA diagnostic model calculated from the HW Raman data yielded a diagnostic sensitivity of 93.5% and specificity of 97.8% for dysplasia tissue, demonstrating the diagnostic ability of HW Raman spectroscopy for cervical dysplasia.

We also made efforts to improve the *in vivo* diagnosis of cervical dysplasia through combining NIR autofluorescence and Raman spectroscopy. Classification results obtained from the leave-one patient-out, cross-validation of the PCA-LDA model based on the three spectral datasets (i.e., Raman, NIR AF, and the combined NIR AF and Raman) yielded diagnostic sensitivities of 91.3%, 93.5%, 93.5%, specificities of 95.7%, 87.0%, 95.7%, and accuracies of 93.5%, 90.2%, 94.6%, respectively, for precancer identification. Obviously, optimal discrimination was achieved by combined NIR AF and Raman, indicating NIR Raman and autofluorescence spectral features are complimentary rather than competitive in tissue diagnosis. The improvement in the diagnosis by combining NIR AF and Raman was further confirmed by ROC testing. The results above demonstrate that the integrated NIR autofluorescence and Raman spectroscopy has the potential to provide an effective and accurate diagnostic scheme for non-invasive, *in vivo* diagnosis and detection of precancer in the cervix.

Alternatively, UV/VIS autofluorescence and diffuse reflectance were also combined with NIR Raman to optimize the optical diagnosis of cervical dysplasia. The diagnostic accuracy yielded by PCA-LDA models indicated that NIR Raman spectroscopy exhibited a better performance in differentiating cervical dysplasia than UV/VIS autofluorescence and diffuse reflectance spectroscopy. The optimal diagnosis was produced by combining Raman, fluorescence and reflectance with maximum rule which means the diagnostic decision is determined by the maximum of the posterior probabilities calculated from Raman, fluorescence and reflectance. The diagnostic accuracies were 100%, 97.1% and 100% for benign, LGSILs and HGSILs, respectively. This demonstrates the complementarity of Raman, fluorescence and reflectance in tissue diagnosis and the feasibility of improving clinic diagnosis with an integration of different optical screen modalities.

7.2 Future Directions

In this study, we demonstrated that NIR Raman spectroscopy has the potential for cervical precancer detection and also exhibit higher diagnostic accuracy as compared conventional screening and diagnosis methods. However, further optimization is still desirable prior to being as a routine clinical screening and diagnosis program to prevent cervical cancer. Specifically, the future works to this thesis study are proposed as follows:

1) Developing a robust spectroscopy program integrated with comprehensive functional modules, including data acquisition, data process and multivariate statistical analysis. The current Raman program built on LabVIEW platform only contains data acquisition and process modules. However, to achieve the true real-time tissue diagnosis and characterization, integration of diagnostic model is necessary. Meanwhile, prior to the on-line diagnosis in clinic, a large database must be built to validate the diagnostic model.

2) Optimizing the diagnostic model by employing other multivariate statistic techniques. In this study, PCA-LDA was used for developing diagnostic algorithms throughout the whole thesis. Apart from PCA and LDA, there exist other multivariate statistic techniques which have been used for developing classification functions, such as support vector machine (SVM), artificial neural network (ANN), cluster analysis, recursive partition and random forest [9, 28, 46, 54, 65, 173, 267, 268]. For example, Widjaja et al used SVM to develop classification model for Raman spectra [28]. A pretty high accuracy of 99.9% was achieved in a multi-classification among normal, polyp and cancer colon tissues. Therefore, a proper selection of multivariate statistic techniques may be one of the choices of method to optimize the diagnostic algorithm

to optimize the diagnosis.

3) Developing an *in vivo* optical diagnostic model integrated with NIR Raman, UV/VIS fluorescence and reflectance spectroscopy. In this thesis, the feasibility of combining these three different optical spectroscopic techniques to improve the diagnosis of cervical dysplasia was demonstrated *ex vivo*. In terms of clinical application, a multi-model suited for *in vivo* spectroscopic measurement is highly desired. Moreover, a higher efficacy in tissue diagnosis and identification may be achieved by adding the imaging mode into the multi-model system in addition to the spectroscopy mode. For instance, imaging mode can be used to rough the malignant tissue quickly and subsequently spectroscopy mode can further confirm the diagnosis and the tumor margin.

4) Building an *in vitro* cervical tissue model to mimic live tissue at various histopathological conditions (i.e., normal, benign, LGSILs and HGSILs) for better understanding the biochemical changes accounting for Raman diagnosis. Although tissue classification is the primary goal of the diagnostic measurements in a clinical setting, understanding the underlying spectral differences is crucial for further validating and optimizing the methodology. In this study, we tentatively attributed the prominent Raman peaks to specific biochemicals and also used mathematical modeling (i.e., least squares fitting) to reconstruct the tissue spectrum with reference spectra of pure biochemicals. However, the accuracy of the former is limited by the overlap of different Raman bands; the latter one may be affected a lot by the selection of reference spectra and also does not consider the interaction among different biochemicals in tissue. Therefore, we propose to use organotypic culture method to build an *in vitro* tissue model to resemble the *in vivo* cervical neoplasia. Organotypic culture has been proven to be able to achieve the stratification and differentiation of

epithelial cells which are histopathologically similar to the epithelium *in vivo* in skin [269, 270], testis [271], ovary [272], stomach [273] and cervix [274, 275]. With this method, we can control and manipulate tissue components and culture conditions to build various tissue models for Raman study. Meanwhile, Raman study on individual epithelial cell lines is proposed as an adjunction to the tissue model study.

Publications

Peer-reviewed Journal Articles

1. **Jianhua Mo**, Wei Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, and Zhiwei Huang, "High wavenumber Raman spectroscopy for in vivo detection of cervical dysplasia," *Analytical Chemistry* **81**(21), 8908 (2009).
2. Zhiwei Huang, Seng Khoon Teh, Wei Zheng, **Jianhua Mo**, Kan Lin, Xiaozhuo Shao, Khek Yu Ho, Ming Teh, and Khay Guan Yeoh, "Integrated Raman spectroscopy and trimodal wide-field imaging techniques for real-time in vivo tissue Raman measurements at endoscopy," *Optics Letters* **34**(6), 758 (2009).
3. **Jianhua Mo**, Wei Zheng, and Zhiwei Huang, "Fiber-optic Raman probe couples ball lens for depth-selected Raman measurements of epithelial tissue," *Biomedical Optics Express* **1**(1), 17 (2010).
4. **Jianhua Mo**, Wei Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, and Zhiwei Huang, "NIR Raman spectroscopy combined with multivariate statistic techniques and spectral modeling for the detection of cervical precancer," in preparation.
5. **Jianhua Mo**, Wei Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, and Zhiwei Huang, "*In vivo* diagnosis of cervical precancer using NIR-excited autofluorescence and Raman spectroscopy," in preparation.
6. **Jianhua Mo**, Wei Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, and Zhiwei Huang, "Evaluating and combining NIR Raman, UV/VIS fluorescence and diffuse reflectance spectroscopy for differentiating cervical dysplasia," in preparation.

Conference Presentations and Proceedings

1. Zhiwei Huang, Effendi Widjaja, Wei Zheng, **Jianhua Mo**, and Colin Sheppard, "Classification of Raman spectra of colonic tissues using pattern recognition technique," in *Biomedical Optics*, Technical Digest (CD) (Optical Society of America 2006), paper number ME62, 19-22 March 2006, Fort Lauderdale, Florida, USA.
2. Zhiwei Huang, Wei Zheng, Effendi Widjaja, **Jianhua Mo**, and Colin Sheppard, "Classification of colonic tissues using Raman spectroscopy and multivariate

techniques," SPIE/BIOS Photonics West: Biomedical Optics, paper number 60930Q, 21-25 January 2006, San Jose, California, USA.

3. **Jianhua Mo**, Wei Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, and Zhiwei Huang, "In vivo diagnosis of cervical precancer using high wavenumber Raman spectroscopy," SPIE/COS Photonics Asia: Optics in Health Care and Biomedical Optics, paper number 682611, 11-55 November 2007, Beijing, China.
4. Zhiwei Huang, **Jianhua Mo**, Wei Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, "Combining near-infrared autofluorescence and Raman spectroscopy improves the in vivo detection of cervical precancer," CLEO/QELS, paper number 1, 4-9 May 2008, San Jose, Carlifornia, USA.
5. **Jianhua Mo**, W. Zheng, Jeffrey J. H. Low, Joseph Ng, A. Ilancheran, and Zhiwei Huang, "Near-infrared autofluorescence spectroscopy for in vivo diagnosis of cervical precancer," SPIE/BIOS Photonics West: Biomedical Optics, paper number 685317, 19-24 January 2008, San Jose, California, USA.
6. Xiaozhuo Shao, **Jianhua Mo**, Wei Zheng, and Zhiwei Huang, "Near-infrared fluorescence imaging for colonic cancer detection," SPIE/BIOS Photonics West: Biomedical Optics, paper number 68480F, 19-24 January 2008, San Jose, California, USA.

References

1. Ferlay, J., F. Bray, P. Pisani, and D.M. Parkin, *BLOBOCAN 2002 Cancer Incidence, Mortality and Prevalence Worldwide.* p. IARC CancerBASE No. 5, version 2.0, IARCPress, Lyon, 2004.
2. Fahey, M.T., L. Irwig, and P. Macaskill, *Meta-analysis of Pap Test Accuracy.* American Journal of Epidemiology, 1995. **141**(7): p. 680-689.
3. Mitchell, M.F., *The Accuracy of Colposcopy.* Clin. Consult. Obst. Gynecol., 1994. **26**: p. 70-73
4. McCreery, R.L., *Raman spectroscopy for chemical analysis*. Chemical analysis. 2000, New York: John Wiley & Sons. 420 p.
5. Mahadevan-Jansen, A. and R.R. Richards-Kortum, *Raman spectroscopy for the detection of cancers and precancers.* Journal of Biomedical Optics, 1996. **1**(1): p. 31-70.
6. Mizuno, A., H. Kitajima, K. Kawauchi, S. Muraishi, and Y. Ozaki, *Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours.* Journal of Raman Spectroscopy, 1994. **25**(1): p. 25-29.
7. Koljenović, S., L.-P. Choo-Smith, T.C. Bakker Schut, J.M. Kros, H.J. van den Berge, and G.J. Puppels, *Discriminating Vital Tumor from Necrotic Tissue in Human Glioblastoma Tissue Samples by Raman Spectroscopy.* Laboratory Investigation, 2002. **82**(10): p. 1265-1277.
8. Koljenović, S., T.C.B. Schut, R. Wolthuis, B.d. Jong, L. Santos, P.J. Caspers, J.M. Kros, and G.J. Puppels, *Tissue characterization using high wave number Raman spectroscopy.* Journal of Biomedical Optics, 2005. **10**(3): p. 031116.
9. Krafft, C., S.B. Sobottka, G. Schackert, and R. Salzer, *Near infrared Raman spectroscopic mapping of native brain tissue and intracranial tumors.* The Analyst, 2005. **130**(7): p. 1070-1077.
10. Köhler, M., S. Machill, R. Salzer, and C. Krafft, *Characterization of lipid extracts from brain tissue and tumors using Raman spectroscopy and mass spectrometry.* Analytical and Bioanalytical Chemistry, 2009. **393**(5): p. 1513-1520.
11. Redd, D.C.B., Z.C. Feng, K.T. Yue, and T.S. Gansler, *Raman Spectroscopic Characterization of Human Breast Tissues: Implications for Breast Cancer Diagnosis.* Applied Spectroscopy, 1993. **47**(6): p. 787-791.
12. Frank, C.J., D.C.B. Redd, T.S. Gansler, and R.L. McCreery, *Characterization of human breast biopsy specimens with near-IR Raman spectroscopy.* Analytical Chemistry, 1994. **66**(3): p. 319-326.
13. Frank, C.J., R.L. McCreery, and D.C.B. Redd, *Raman Spectroscopy of Normal and Diseased Human Breast Tissues.* Analytical Chemistry, 1995. **67**(5): p. 777-783.
14. Manoharan, R., K. Shafer, L. Perelman, J. Wu, K. Chen, G. Deinum, M. Fitzmaurice, J. Myles, J. Crowe, R.R. Dasarl, and M.S. Feld, *Raman Spectroscopy and Fluorescence Photon Migration for Breast Cancer Diagnosis and Imaging.* Photochemistry and Photobiology, 1998. **67**(1): p. 15-22.
15. Chowdary, M.V.P., K.K. Kumar, J. Kurien, S. Mathew, and C.M. Krishna, *Discrimination of normal, benign, and malignant breast tissues by Raman spectroscopy.* Biopolymers, 2006. **83**(5): p. 556-569.
16. Shafer-Peltier, K.E., A.S. Haka, M. Fitzmaurice, J. Crowe, J. Myles, R.R.

Dasari, and M.S. Feld, *Raman microspectroscopic model of human breast tissue: implications for breast cancer diagnosis in vivo.* Journal of Raman Spectroscopy, 2002. **33**(7): p. 552-563.
17. Haka, A.S., K.E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R.R. Dasari, and M.S. Feld, *Diagnosing breast cancer by using Raman spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(35): p. 12371-12376.
18. Haka, A.S., Z. Volynskaya, J.A. Gardecki, J. Nazemi, J. Lyons, D. Hicks, M. Fitzmaurice, R.R. Dasari, J.P. Crowe, and M.S. Feld, *In vivo Margin Assessment during Partial Mastectomy Breast Surgery Using Raman Spectroscopy[?Q1: Running head: Raman Margin Assessment at Partial Mastectomy. Short title OK?Q1]*. Cancer Research, 2006. **66**(6): p. 3317-3322.
19. Haka, A.S., Z. Volynskaya, J.A. Gardecki, J. Nazemi, R. Shenk, N. Wang, R.R. Dasari, M. Fitzmaurice, and M.S. Feld, *Diagnosing breast cancer using Raman spectroscopy: prospective analysis.* Journal of Biomedical Optics, 2009. **14**(5): p. 054023-8.
20. Chowdary, M.V.P., K.K. Kumar, S. Mathew, L. Rao, C.M. Krishna, and J. Kurien, *Biochemical correlation of Raman spectra of normal, benign and malignant breast tissues: A spectral deconvolution study.* Biopolymers, 2009. **91**(7): p. 539-546.
21. Haka, A.S., K.E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R.R. Dasari, and M.S. Feld, *Identifying Microcalcifications in Benign and Malignant Breast Lesions by Probing Differences in Their Chemical Composition Using Raman Spectroscopy.* Cancer Research, 2002. **62**(18): p. 5375-5380.
22. Matousek, P. and N. Stone, *Prospects for the diagnosis of breast cancer by noninvasive probing of calcifications using transmission Raman spectroscopy.* Journal of Biomedical Optics, 2007. **12**(2): p. 024008-8.
23. Keller, S., B. Schrader, A. Hoffmann, W. Schrader, K. Metz, A. Rehlaender, J. Pahnke, M. Ruwe, and W. Budach, *Application of near-infrared-Fourier transform Raman spectroscopy in medical research.* Journal of Raman Spectroscopy, 1994. **25**(7-8): p. 663-671.
24. Feld, M.S., 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*. in *Advances in Fluorescence Sensing Technology II*. 1995. San Jose, CA, USA: SPIE.
25. Shim, M.G., L.M. Song, N.E. Marcon, and B.C. Wilson, *In vivo near-infrared Raman spectroscopy: demonstration of feasibility during clinical gastrointestinal endoscopy &para.* Photochemistry and Photobiology, 2000. **72**(1): p. 146-150.
26. Molckovsky, A., L.-M.W.K. Song, M.G. Shim, N.E. Marcon, and B.C. Wilson, *Diagnostic potential of near-infrared Raman spectroscopy in the colon: Differentiating adenomatous from hyperplastic polyps.* Gastrointestinal Endoscopy, 2003. **57**(3): p. 396-402.
27. Andrade, P., R. Bitar, K. Yassoyama, H. Martinho, A. Santo, P. Bruno, and A. Martin, *Study of normal colorectal tissue by FT-Raman spectroscopy.* Analytical and Bioanalytical Chemistry, 2007. **387**(5): p. 1643-1648.
28. Widjaja, E., W. Zheng, and Z. Huang, *Classification of colonic tissues using near-infrared Raman spectroscopy and support vector machines.* International Journal of Oncology, 2008. **32**: p. 653-662.
29. Song, L.-M.W.K., A. Molckovsky, K.K. Wang, L.J. Burgart, B. Dolenko, R.L.

Somorjai, and B.C. Wilson. *Diagnostic potential of Raman spectroscopy in Barrett's esophagus*. in *Advanced Biomedical and Clinical Diagnostic Systems III*. 2005. San Jose, CA, USA: SPIE.
30. Kendall, C., 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.* The Journal of Pathology, 2003. **200**(5): p. 602-609.
31. Shetty, G., C. Kendall, N. Shepherd, N. Stone, and H. Barr, *Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus.* British Journal of Cancer, 2006. **94**(10): p. 1460-1464.
32. Hattori, Y., Y. Komachi, T. Asakura, T. Shimosegawa, G.-I. Kanai, H. Tashiro, and H. Sato, *In Vivo Raman Study of the Living Rat Esophagus and Stomach Using a Micro-Raman Probe Under an Endoscope.* Applied Spectroscopy, 2007. **61**(6): p. 579-584.
33. Day, J.C.C., R. Bennett, B. Smith, C. Kendall, J. Hutchings, G.M. Meaden, C. Born, S. Yu, and N. Stone, *A miniature confocal Raman probe for endoscopic use.* Physics in Medicine and Biology, 2009. **54**(23): p. 7077.
34. Nie, S., D.C.B. Redd, Y. Li, and N.-T. Yu. *Near-IR Fourier transform Raman spectroscopy in surgery and medicine: detection of renal stones and bladder cancer*. in *Laser Surgery: Advanced Characterization, Therapeutics, and Systems III*. 1992. Los Angeles, CA, USA: SPIE.
35. de Jong, B.W.D., T.C. Bakker Schut, K.P. Wolffenbuttel, J.M. Nijman, D.J. Kok, and G.J. Puppels, *Identification of Bladder Wall Layers by Raman Spectroscopy.* The Journal of Urology, 2002. **168**(4, Supplement 1): p. 1771-1778.
36. de Jong, B.W.D., T.C.B. Schut, J. Coppens, K.P. Wolffenbuttel, D.J. Kok, and G.J. Puppels, *Raman spectroscopic detection of changes in molecular composition of bladder muscle tissue caused by outlet obstruction.* Vibrational Spectroscopy, 2003. **32**(1): p. 57-65.
37. Crow, P., N. Stone, C.A. Kendall, J.S. Uff, J.A.M. Farmer, H. Barr, and M.P.J. Wright, *The use of Raman spectroscopy to identify and grade prostatic adenocarcinoma in vitro.* British Journal of Cancer, 2003. **89**(1): p. 106-108.
38. Crow, P., J.S. Uff, J.A. Farmer, M.P. Wright, and N. Stone, *The use of Raman spectroscopy to identify and characterize transitional cell carcinoma in vitro.* BJU International, 2004. **93**(9): p. 1232-1236.
39. Crow, P., A. Molckovsky, N. Stone, J. Uff, B. Wilson, and L.M.W.K. Song, *Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer.* Urology, 2005. **65**(6): p. 1126-1130.
40. Crow, P., B. Barrass, C. Kendall, M. Hart-Prieto, M. Wright, R. Persad, and N. Stone, *The use of Raman spectroscopy to differentiate between different prostatic adenocarcinoma cell lines.* British Journal of Cancer, 2005. **92**(12): p. 2166-2170.
41. Taleb, A., J. Diamond, J.J. McGarvey, J.R. Beattie, C. Toland, and P.W. Hamilton, *Raman Microscopy for the Chemometric Analysis of Tumor Cells.* The Journal of Physical Chemistry B, 2006. **110**(39): p. 19625-19631.
42. de Jong, B.W.D., T.C. Bakker Schut, K. Maquelin, T. van der Kwast, C.H. Bangma, D.-J. Kok, and G.J. Puppels, *Discrimination between Nontumor Bladder Tissue and Tumor by Raman Spectroscopy.* Analytical Chemistry, 2006. **78**(22): p. 7761-7769.
43. Stone, N., M. Hart Prieto, P. Crow, J. Uff, and A. Ritchie, *The use of Raman spectroscopy to provide an estimation of the gross biochemistry associated with*

urological pathologies. Analytical and Bioanalytical Chemistry, 2007. **387**(5): p. 1657-1668.
44. Stone, N., P. Stavroulaki, C. Kendall, M. Birchall, and H. Barr, *Raman Spectroscopy for Early Detection of Laryngeal Malignancy: Preliminary Results.* The Laryngoscope, 2000. **110**(10): p. 1756-1763.
45. Lau, D.P., Z. Huang, H. Lui, D.W. Anderson, K. Berean, M.D. Morrison, L. Shen, and H. Zeng, *Raman spectroscopy for optical diagnosis in the larynx: Preliminary findings.* Lasers in Surgery and Medicine, 2005. **37**(3): p. 192-200.
46. Teh, S.K., W. Zheng, D.P. Lau, and Z. Huang, *Spectroscopic diagnosis of laryngeal carcinoma using near-infrared Raman spectroscopy and random recursive partitioning ensemble techniques.* The Analyst, 2009. **134**(6): p. 1232-1239.
47. Feng, S., J. Lin, M. Cheng, Y.-Z. Li, G. Chen, Z. Huang, Y. Yu, R. Chen, and H. Zeng, *Gold Nanoparticle Based Surface-Enhanced Raman Scattering Spectroscopy of Cancerous and Normal Nasopharyngeal Tissues Under Near-Infrared Laser Excitation.* Applied Spectroscopy, 2009. **63**(10): p. 1089-1094.
48. Buiteveld, H., F.F.M. de Mul, J. Mud, and J. Greve, *Identification of Inclusions in Lung Tissue with a Raman Microprobe.* Applied Spectroscopy, 1984. **38**(3): p. 304-306.
49. Schut, T.C.B., G.J. Puppels, Y.M. Kraan, J. Greve, L.L.J.v.d. Maas, and C.G. Figdor, *Intracellular carotenoid levels measured by Raman microspectroscopy: Comparison of lymphocytes from lung cancer patients and healthy individuals.* International Journal of Cancer, 1997. **74**(1): p. 20-25.
50. Kaminaka, S., H. Yamazaki, T. Ito, E. Kohda, and H.-o. Hamaguchi, *Near-infrared Raman spectroscopy of human lung tissues: possibility of molecular-level cancer diagnosis.* Journal of Raman Spectroscopy, 2001. **32**(2): p. 139-141.
51. Huang, Z., A. McWilliams, H. Lui, D.I. McLean, S. Lam, and H. Zeng, *Near-infrared Raman spectroscopy for optical diagnosis of lung cancer.* International Journal of Cancer, 2003. **107**(6): p. 1047-1052.
52. Huang, Z., A. McWilliams, S. Lam, J. English, D.I. McLean, H. Lui, and H. Zeng, *Effect of formalin fixation on the near-infrared Raman spectroscopy of normal and cancerous human bronchial tissues.* International Journal of Oncology, 2003. **23**: p. 649-655.
53. Short, M.A., S. Lam, A. McWilliams, J. Zhao, H. Lui, and H. Zeng, *Development and preliminary results of an endoscopic Raman probe for potential in vivo diagnosis of lung cancers.* Optics Letters, 2008. **33**(7): p. 711-713.
54. Koljenovic, S., T.C.B. Schut, J.P. van Meerbeeck, A.P.W.M. Maat, S.A. Burgers, P.E. Zondervan, J.M. Kros, and G.J. Puppels, *Raman microspectroscopic mapping studies of human bronchial tissue.* Journal of Biomedical Optics, 2004. **9**(6): p. 1187-1197.
55. Krafft, C., D. Codrich, G. Pelizzo, and V. Sergo, *Raman mapping and FTIR imaging of lung tissue: congenital cystic adenomatoid malformation.* The Analyst, 2008. **133**(3): p. 361-371.
56. Jess, P.R.T., M. Mazilu, K. Dholakia, A.C. Riches, and C.S. Herrington, *Optical detection and grading of lung neoplasia by Raman microspectroscopy.* International Journal of Cancer, 2009. **124**(2): p. 376-380.
57. Oshima, Y., H. Shinzawa, T. Takenaka, C. Furihata, and H. Sato,

Discrimination analysis of human lung cancer cells associated with histological type and malignancy using Raman spectroscopy. Journal of Biomedical Optics, 2010. **15**(1): p. 017009-8.

58. Gniadecka, M., H.C. Wulf, N.N. Mortensen, O.F. Nielsen, and D.H. Christensen, *Diagnosis of Basal Cell Carcinoma by Raman Spectroscopy.* Journal of Raman Spectroscopy, 1997. **28**(2-3): p. 125-129.
59. Nijssen, A., T.C. Bakker 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.* 2002. **119**(1): p. 64-69.
60. Nijssen, A., K. Maquelin, L.F. Santos, P.J. Caspers, T.C.B. Schut, J.C.d. Hollander, M.H.A. Neumann, and G.J. Puppels, *Discriminating basal cell carcinoma from perilesional skin using high wave-number Raman spectroscopy.* Journal of Biomedical Optics, 2007. **12**(3): p. 034004.
61. Short, M.A., H. Lui, D. McLean, H. Zeng, A. Alajlan, and X.K. Chen, *Changes in nuclei and peritumoral collagen within nodular basal cell carcinomas via confocal micro-Raman spectroscopy.* Journal of Biomedical Optics, 2006. **11**(3): p. 034004-9.
62. Ly, E., O. Piot, A. Durlach, P. Bernard, and M. Manfait, *Polarized Raman Microspectroscopy Can Reveal Structural Changes of Peritumoral Dermis in Basal Cell Carcinoma.* Applied Spectroscopy, 2008. **62**(10): p. 1088-1094.
63. Larraona-Puy, M., A. Ghita, A. Zoladek, W. Perkins, S. Varma, I.H. Leach, A.A. Koloydenko, H. Williams, and I. Notingher, *Development of Raman microspectroscopy for automated detection and imaging of basal cell carcinoma.* Journal of Biomedical Optics, 2009. **14**(5): p. 054031-10.
64. Gniadecka, M., H.C. Wulf, O.F. Nielsen, D.H. Christensen, and J. Hercogova, *Distinctive Molecular Abnormalities in Benign and Malignant Skin Lesions: Studies by Raman Spectroscopy.* Photochemistry and Photobiology, 1997. **66**(4): p. 418-423.
65. Gniadecka, M., P.A. Philipsen, S. Sigurdsson, S. Wessel, O.F. Nielsen, D.H. Christensen, J. Hercogova, K. Rossen, H.K. Thomsen, R. Gniadecki, L.K. Hansen, and H.C. Wulf, *Melanoma Diagnosis by Raman Spectroscopy and Neural Networks: Structure Alterations in Proteins and Lipids in Intact Cancer Tissue.* J Investig Dermatol, 2004. **122**(2): p. 443-449.
66. Lieber, C.A., S.K. Majumder, D. Billheimer, D.L. Ellis, and A. Mahadevan-Jansen, *Raman microspectroscopy for skin cancer detection in vitro.* Journal of Biomedical Optics, 2008. **13**(2): p. 024013-9.
67. Huang, Z., H. Zeng, I. Hamzavi, D.I. McLean, and H. Lui, *Rapid near-infrared Raman spectroscopy system for real-time in vivo skin measurements.* Optics Letters, 2001. **26**(22): p. 1782-1784.
68. Lieber, C. and A. Mahadevan-Jansen, *Development of a handheld Raman microspectrometer for clinical dermatologic applications.* Optics Express, 2007. **15**(19): p. 11874-11882.
69. Lieber, C.A., S.K. Majumder, D.L. Ellis, D.D. Billheimer, and A. Mahadevan-Jansen, *In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy.* Lasers in Surgery and Medicine, 2008. **40**(7): p. 461-467.
70. Teh, S.K., W. Zheng, K.Y. Ho, M. Teh, K.G. Yeoh, and Z. Huang, *Diagnostic potential of near-infrared Raman spectroscopy in the stomach: differentiating dysplasia from normal tissue.* British Journal of Cancer, 2008. **98**(2): p. 457-465.
71. Teh, S.K., W. Zheng, K.Y. Ho, M. Teh, K.G. Yeoh, and Z. Huang, *Near-infrared*

Raman spectroscopy for gastric precancer diagnosis. Journal of Raman Spectroscopy, 2009. **40**(8): p. 908-914.
72. Teh, S.K., W. Zheng, K.Y. Ho, M. Teh, K.G. Yeoh, and Z. Huang, *Diagnosis of gastric cancer using near-infrared Raman spectroscopy and classification and regression tree techniques.* Journal of Biomedical Optics, 2008. **13**(3): p. 034013-8.
73. Teh, S.K., W. Zheng, K.Y. Ho, M. Teh, K.G. Yeoh, and Z. Huang, *Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach.* British Journal of Surgery 2010. **97**: p. 550-557.
74. Teh, S.K., W. Zheng, K.Y. Ho, M. Teh, K.G. Yeoh, and Z. Huang, *Near-infrared Raman spectroscopy for optical diagnosis in the stomach: Identification of Helicobacter-pylori infection and intestinal metaplasia.* International Journal of Cancer, 2010. **126**(8): p. 1920-1927.
75. Kalyan Kumar, K., A. Anand, M.V.P. Chowdary, Keerthi, J. Kurien, C. Murali Krishna, and S. Mathew, *Discrimination of normal and malignant stomach mucosal tissues by Raman spectroscopy: A pilot study.* Vibrational Spectroscopy, 2007. **44**(2): p. 382-387.
76. Hu, Y., A. Shen, T. Jiang, Y. Ai, and J. Hu, *Classification of normal and malignant human gastric mucosa tissue with confocal Raman microspectroscopy and wavelet analysis.* Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2008. **69**(2): p. 378-382.
77. Kawabata, T., T. Mizuno, S. Okazaki, M. Hiramatsu, T. Setoguchi, H. Kikuchi, M. Yamamoto, Y. Hiramatsu, K. Kondo, M. Baba, M. Ohta, K. Kamiya, T. Tanaka, S. Suzuki, and H. Konno, *Optical diagnosis of gastric cancer using near-infrared multichannel Raman spectroscopy with a 1064-nm excitation wavelength.* Journal of Gastroenterology, 2008. **43**(4): p. 283-290.
78. Huang, Z., S.K. Teh, W. Zheng, J. Mo, K. Lin, X. Shao, K.Y. Ho, M. Teh, and K.G. Yeoh, *Integrated Raman spectroscopy and trimodal wide-field imaging techniques for real-time in vivo tissue Raman measurements at endoscopy.* Optics Letters, 2009. **34**(6): p. 758-760.
79. Notani, P.N., *Global variation in cancer incidence and mortality.* Current Science, 2001. **81**(5): p. 465-474.
80. Yang, B.H., F.I. Bray, D.M. Parkin, J.W. Sellors, and Z.-F. Zhang, *Cervical cancer as a priority for prevention in different world regions: An evaluation using years of life lost.* International Journal of Cancer, 2004. **109**(3): p. 418-424.
81. Clifford, G., S. Franceschi, M. Diaz, N. Muñoz, and L.L. Villa, *Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases.* Vaccine, 2006. **24**(Supplement 3): p. S26-S34.
82. Shields, T.S., L.A. Brinton, R.D. Burk, S.S. Wang, S.J. Weinstein, R.G. Ziegler, Y.Y. Studentsov, M. McAdams, and M. Schiffman, *A Case-Control Study of Risk Factors for Invasive Cervical Cancer among U.S. Women Exposed to Oncogenic Types of Human Papillomavirus.* Cancer Epidemiology Biomarkers & Prevention, 2004. **13**(10): p. 1574-1582.
83. Moreno, V., F.X. Bosch, N. Muñoz, C.J.L.M. Meijer, K.V. Shah, J.M.M. Walboomers, R. Herrero, and S. Franceschi, *Effect of oral contraceptives on risk of cervical cancer in women with human papillomavirus infection: the IARC multicentric case-control study.* The Lancet, 2002. **359**(9312): p. 1085-1092.
84. Monsonego, J., F.X. Bosch, P. Coursaget, J.T. Cox, E. Franco, I. Frazer, R.

Sankaranarayanan, J. Schiller, A. Singer, T. Wright, W. Kinney, C. Meijer, and J. Linder, *Cervical cancer control, priorities and new directions.* International Journal of Cancer, 2004. **108**(3): p. 329-333.
85. Hemminki, K., C. Dong, and P. Vaittinen, *Familial risks in cervical cancer: Is there a hereditary component?* International Journal of Cancer, 1999. **82**(6): p. 775-781.
86. Magnusson, P.K.E., P. Sparén, and U.B. Gyllensten, *Genetic link to cervical tumours.* Nature, 1999. **400**(6739): p. 29-30.
87. Kjellberg, L., G. Hallmans, A.M. Ahren, R. Johansson, F. Bergman, G. Wadell, T. Angstrom, and J. Dillner, *Smoking, diet, pregnancy and oral contraceptive use as risk factors for cervical intra-epithelial neoplasia in relation to human papillomavirus infection.* British Journal of Cancer, 2000. **82**(7): p. 1332-1338.
88. Parazzini, F., C.L. Vecchia, E. Negri, L. Fedele, S. Franceschi, and L. Gallotta, *Risk factors for cervical intraepithelial neoplasia.* Cancer, 1992. **69**(9): p. 2276-2282.
89. Kumar, V., A.K. Abbas, N. Fausto, S.L. Robbins, and R.S. Cotran. *Robbins and Cotran pathologic basis of disease*. 2005; 7th:[xv, 1525 p.].
90. Buckley, C.H., E.B. Butler, and H. Fox, *Cervical intraepithelial neoplasia.* Journal of Clinical Pathology, 1982. **35**(1): p. 1-13.
91. Richart, R.M., *Natural history of cervical intraepithelial neoplasia.* Clinical Obstetrics and Gynecology, 1967. **10**: p. 748-784.
92. Ismail, S.M., A.B. Colclough, J.S. Dinnen, D. Eakins, D.M. Evans, E. Gradwell, J.P. O'Sullivan, J.M. Summerell, and R.G. Newcombe, *Observer variation in histopathological diagnosis and grading of cervical intraepithelial neoplasia.* BMJ, 1989. **298**(6675): p. 707-710.
93. Diane, S., *The Bethesda System for Reporting Cervical/Vaginal Cytologic Diagnoses:An Overview.* International Journal of Gynecological Pathology, 1991. **10**: p. 323-325.
94. Fletcher, C.D.M., *Diagnostic histopathology of tumors*. 3rd ed. 2007, [Edinburgh] ; Philadelphia: Churchill Livingstone Elsevier. 2 v. (x, 1883, 41 p.).
95. Schiffman, M., P.E. Castle, J. Jeronimo, A.C. Rodriguez, and S. Wacholder, *Human papillomavirus and cervical cancer.* The Lancet. **370**(9590): p. 890-907.
96. Long, H.J., N.N.I. Laack, and B.S. Gostout, *Prevention, Diagnosis, and Treatment of Cervical Cancer.* Mayo Clinic Proceedings, 2007. **82**(12): p. 1566-1574.
97. Liu, C.H., B.B. Das, W.L.S. Glassman, G.C. Tang, K.M. Yoo, H.R. Zhu, D.L. Akins, S.S. Lubicz, J. Cleary, R. Prudente, E. Celmer, A. Caron, and R.R. Alfano, *Raman, fluorescence, and time-resolved light scattering as optical diagnostic techniques to separate diseased and normal biomedical media.* Journal of Photochemistry and Photobiology B: Biology, 1992. **16**(2): p. 187-209.
98. Mahadevan-Jansen, A., M.F. Mitchell, N. Ramanujamf, A. Malpica, S. Thomsen, U. Utzinger, and R. Richards-Kortumt, *Near-infrared Raman spectroscopy for in vitro detection of cervical precancers.* Photochemistry and Photobiology, 1998. **68**(1): p. 123-132.
99. Mahadevan-Jansen, A., M.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, 1998.

68(3): p. 427-431.
100. Utzinger, U., 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, 2001. **55**: p. 955-959.
101. Robichaux-Viehoever, A., 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, 2007. **61**(9): p. 986-993.
102. Kanter, E.M., S. Majumder, E. Vargis, A. Robichaux-Viehoever, G.J. Kanter, H. Shappell, H.W. Jones III, and A. Mahadevan-Jansen, *Multiclass discrimination of cervical precancers using Raman spectroscopy.* Journal of Raman Spectroscopy, 2009. **40**(2): p. 205-211.
103. Kanter, E.M., S. Majumder, G.J. Kanter, E.M. Woeste, and A. Mahadevan-Jansen, *Effect of hormonal variation on Raman spectra for cervical disease detection.* American Journal of Obstetrics and Gynecology, 2009. **200**(5): p. 511-515.
104. Jess, P.R.T., D.D.W. Smith, M. Mazilu, K. Dholakia, A.C. Riches, and C.S. Herrington, *Early detection of cervical neoplasia by Raman spectroscopy.* International Journal of Cancer, 2007. **121**(12): p. 2723-2728.
105. Martinho, H.d.S., C.M.d.O.M.d. Silva, M.C.B.M. Yassoyama, P.d.O. Andrade, R.A. Bitar, A.M.d.E. Santo, E.A.L. Arisawa, and A.A. Martin, *Role of cervicitis in the Raman-based optical diagnosis of cervical intraepithelial neoplasia.* Journal of Biomedical Optics, 2008. **13**(5): p. 054029.
106. Krishna, C.M., N.B. Prathima, R. Malini, B.M. Vadhiraja, R.A. Bhatt, D.J. Fernandes, P. Kushtagi, M.S. Vidyasagar, and V.B. Kartha, *Raman spectroscopy studies for diagnosis of cancers in human uterine cervix.* Vibrational Spectroscopy, 2006. **41**(1): p. 136-141.
107. Krishna, C.M., G.D. Sockalingum, B.M. Vadhiraja, K. Maheedhar, A.C.K. Rao, L. Rao, L. Venteo, M. Pluot, D.J. Fernandes, M.S. Vidyasagar, V.B. Kartha, and M. Manfait, *Vibrational spectroscopy studies of formalin-fixed cervix tissues.* Biopolymers, 2007. **85**(3): p. 214-221.
108. Lyng, F.M., E.Ó. Faoláin, J. Conroy, A.D. Meade, P. Knief, B. Duffy, M.B. Hunter, J.M. Byrne, P. Kelehan, and H.J. Byme, *Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool.* Experimental and Molecular Pathology, 2007. **82**(2): p. 121-129.
109. Kamemoto, L.E., A.K. Misra, S.K. Sharma, M.T. Goodman, H. Luk, A.C. Dykes, and T. Acosta, *Near-Infrared Micro-Raman Spectroscopy for in Vitro Detection of Cervical Cancer.* Applied Spectroscopy, 2010. **64**(3): p. 255-261.
110. Lohmann, W., J. Mumann, C. Lohmann, and W. Küzel, *Native fluorescence of the cervix uteri as a marker for dysplasia and invasive carcinoma.* European Journal of Obstetrics & Gynecology and Reproductive Biology, 1989. **31**(3): p. 249-253.
111. Lohmann, W., J. Mussmann, C. Lohmann, and W. Künzel, *Native fluorescence of unstained cryo-sections of the cervix uteri compared with histological observations.* Naturwissenschaften, 1989. **76**(3): p. 125-127.
112. Glassman, W.S., C.H. Liu, G.C. Tang, S. Lubicz, and R.R. Alfano, *Ultraviolet excited fluorescence spectra from non-malignant and malignant tissues of the gynecological tract.* Lasers in the Life Sciences, 1992. **5**(1-2): p. 49-58.
113. Glassman, W.S., C.H. Liu, G.C. Tang, S. Lubicz, and R.R. Alfano, *Excitation spectroscopy of malignant and non-malignant gynecological tissues.* Lasers in

the Life Sciences, 1994. **6**(2): p. 99-106.
114. Mahadevan, A., M.F. Mitchell, E. Silva, S. Thomsen, and R.R. Richards-Kortum, *Study of the fluorescence properties of normal and neoplastic human cervical tissue.* Lasers in Surgery and Medicine, 1993. **13**(6): p. 647-655.
115. Koumantakis, E., A. Vasileiou, A. Makrigiannakis, E. Unsöld, and T.G. Papazoglou, *Spectral variations of laser-induced tissue emission during in vivo detection of malignancies in the female genital tract.* Journal of Photochemistry and Photobiology B: Biology, 1997. **40**(2): p. 183-186.
116. Ramanujam, N., M.F. Mitchell, A. Mahadevan, S. Thomsen, E. Silva, and R. Richards-Kortum, *Fluorescence Spectroscopy: A Diagnostic Tool for Cervical Intraepithelial Neoplasia (CIN).* Gynecologic Oncology, 1994. **52**(1): p. 31-38.
117. Ramanujam, N., M.F. Mitchell, A. Mahadevan, S. Warren, S. Thomsen, E. Silva, and R. Richards-Kortum, *In vivo diagnosis of cervical intraepithelial neoplasia using 337-nm-excited laser-induced fluorescence.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(21): p. 10193-10197.
118. Ramanujam, N., M.F. Mitchell, A. Mahadevan, S. Thomsen, A. Malpica, T. Wright, N. Atkinson, and R. Richards-Kortum, *Development of a multivariate statistical algorithm to analyze human cervical tissue fluorescence spectra acquired in vivo.* Lasers in Surgery and Medicine, 1996. **19**(1): p. 46-62.
119. Ramanujam, N., M.F. Mitchell, A. Mahadevan, S. Thomsen, A. Malpica, T. Wright, N. Atkinson, and R.R. Kortum, *Spectroscopic diagnosis of cervical intraepithelial neoplasia (CIN) in vivo using laser-induced fluorescence spectra at multiple excitation wavlengths.* Lasers in Surgery and Medicine, 1996. **19**(1): p. 63-74.
120. Utzinger, U., 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, 1999. **46**(11): p. 1293-1303.
121. Agrawal, A., U. Utzinger, C. Brookner, C. Pitris, M.F. Mitchell, and R. Richards-Kortum, *Fluorescence spectroscopy of the cervix: Influence of acetic acid, cervical mucus, and vaginal medications.* Lasers in Surgery and Medicine, 1999. **25**(3): p. 237-249.
122. Chang, S.K., M.Y. Dawood, G. Staerkel, U. Utzinger, E.N. Atkinson, R.R. Richards-Kortum, and M. Follen, *Fluorescence spectroscopy for cervical precancer detection: Is there variance across the menstrual cycle?* Journal of Biomedical Optics, 2002. **7**(4): p. 595-602.
123. Cox, D.D., 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, 2003. **57**(1): p. 67-72.
124. Gill, E.M., 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, 2003. **77**(6): p. 653-658.
125. Atkinson, E.N., *Age and FSH effects in fluorescence spectra from the cervix: An exploratory analysis.* Gynecologic Oncology, 2005. **99**(3, Supplement 1): p. S95-S97.

126. Brookner, C.K., U. Utzinger, G. Staerkel, R. Richards-Kortum, and M.F. Mitchell, *Cervical fluorescence of normal women.* Lasers in Surgery and Medicine, 1999. **24**(1): p. 29-37.
127. Nath, A., K. Rivoire, S. Chang, D. Cox, E.N. Atkinson, M. Follen, and R. Richards-Kortum, *Effect of probe pressure on cervical fluorescence spectroscopy measurements.* Journal of Biomedical Optics, 2004. **9**(3): p. 523-533.
128. Rivoire, K., A. Nath, D. Cox, E.N. Atkinson, R. Richards-Kortum, and M. Follen, *The effects of repeated spectroscopic pressure measurements on fluorescence intensity in the cervix.* American Journal of Obstetrics and Gynecology, 2004. **191**(5): p. 1606-1617.
129. Freeberg, J.A., D.M. Serachitopol, N. McKinnon, R. Price, E.N. Atkinson, D.D. Cox, C. MacAulay, R. Richards-Kortum, M. Follen, and B. Pikkula, *Fluorescence and reflectance device variability throughout the progression of a phase II clinical trial to detect and screen for cervical neoplasia using a fiber optic probe.* Journal of Biomedical Optics, 2007. **12**(3): p. 034015-11.
130. Pikkula, B.M., O. Shuhatovich, R.L. Price, D.M. Serachitopol, M. Follen, N. McKinnon, C. MacAulay, R. Richards-Kortum, J.S. Lee, E.N. Atkinson, and D.D. Cox, *Instrumentation as a source of variability in the application of fluorescence spectroscopic devices for detecting cervical neoplasia.* Journal of Biomedical Optics, 2007. **12**(3): p. 034014-11.
131. Brookner, C.K., M. Follen, I. Boiko, J. Galvan, S. Thomsen, A. Malpica, S. Suzuki, R. Lotan, and R. Richards-Kortum, *Autofluorescence Patterns in Short-Term Cultures of Normal Cervical Tissue.* Photochemistry and Photobiology, 2000. **71**(6): p. 730-736.
132. Ramanujam, N., R. Richards-Kortum, S. Thomsen, A. Mahadevan-Jansen, M. Follen, and B. Chance, *Low Temperature Fluorescence Imaging of Freeze-trapped Human Cervical Tissues.* Optics Express, 2001. **8**(6): p. 335-343.
133. Drezek, R., C. Brookner, I. Pavlova, I. Boiko, A. Malpica, R. Lotan, M. Follen, and R. Richards-Kortum, *Autofluorescence microscopy of fresh cervical-tissue sections reveals alterations in tissue biochemistry with dysplasia.* Photochemistry and Photobiology, 2001. **73**(6): p. 636-641.
134. Drezek, R., K. Sokolov, U. Utzinger, I. Boiko, A. Malpica, M. Follen, and R. Richards-Kortum, *Understanding the contributions of NADH and collagen to cervical tissue fluorescence spectra: Modeling, measurements, and implications.* Journal of Biomedical Optics, 2001. **6**(4): p. 385-396.
135. Pavlova, I., K. Sokolov, R. Drezek, A. Malpica, M. Follen, and R. Richards-Kortum, *Microanatomical and Biochemical Origins of Normal and Precancerous Cervical Autofluorescence Using Laser-scanning Fluorescence Confocal Microscopy &para.* Photochemistry and Photobiology, 2003. **77**(5): p. 550-555.
136. Carlson, K., I. Pavlova, T. Collier, M. Descour, M. Follen, and R. Richards-Kortum, *Confocal microscopy: Imaging cervical precancerous lesions.* Gynecologic Oncology, 2005. **99**(3, Supplement 1): p. S84-S88.
137. Wu, Y., P. Xi, J.Y. Qu, T.-H. Cheung, and M.-Y. Yu, *Depth-resolved fluorescence spectroscopy of normal and dysplastic cervical tissue.* Optics Express, 2005. **13**: p. 382-388
138. Georgakoudi, I., B.C. Jacobson, M.G. Muller, E.E. Sheets, K. Badizadegan, D.L. Carr-Locke, C.P. Crum, C.W. Boone, R.R. Dasari, J. Van Dam, and M.S.

Feld, *NAD(P)H and collagen as in vivo quantitative fluorescent biomarkers of epithelial precancerous changes.* Cancer Research, 2002. **62**(3): p. 682-687.
139. Chang, S.K., N. Marin, M. Follen, and R. Richards-Kortum, *Model-based analysis of clinical fluorescence spectroscopy for in vivo detection of cervical intraepithelial dysplasia.* Journal of Biomedical Optics, 2006. **11**(2): p. 024008-12.
140. Chang, S.K., M. Follen, A. Malpica, U. Utzinger, G. Staerkel, D. Cox, E. Neely Atkinson, C. MacAulay, and R. Richards-Kortum, *Optimal excitation wavelengths for discrimination of cervical neoplasia.* Biomedical Engineering, IEEE Transactions on, 2002. **49**(10): p. 1102-1111.
141. Benavides, J., S. Chang, S. Park, R. Richards-Kortum, N. Mackinnon, C. MacAulay, A. Milbourne, A. Malpica, and M. Follen, *Multispectral digital colposcopy for in vivo detection of cervical cancer.* Optics Express, 2003. **11**(10): p. 1223-1236.
142. Milbourne, A., 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.* Gynecologic Oncology, 2005. **99**(3, Supplement 1): p. S67-S75.
143. Heintzelman, D.L., R. Lotan, and R.R. Richards-Kortum, *Characterization of the Autofluorescence of Polymorphonuclear Leukocytes, Mononuclear Leukocytes and Cervical Epithelial Cancer Cells for Improved Spectroscopic Discrimination of Inflammation from Dysplasia.* Photochemistry and Photobiology, 2000. **71**(3): p. 327-332.
144. Nordstrom, R.J., L. Burke, J.M. Niloff, and J.F. Myrtle, *Identification of cervical intraepithelial neoplasia (CIN) using UV-excited fluorescence and diffuse-reflectance tissue spectroscopy.* Lasers in Surgery and Medicine, 2001. **29**(2): p. 118-127.
145. Weingandt, H., H. Stepp, R. Baumgartner, J. Diebold, W. Xiang, and P. Hillemanns, *Autofluorescence spectroscopy for the diagnosis of cervical intraepithelial neoplasia.* BJOG: An International Journal of Obstetrics & Gynaecology, 2002. **109**(8): p. 947-951.
146. Rodero, A., L. Silveira Jr, D. Rodero, R. Racanicchi, and M. Pacheco, *Fluorescence Spectroscopy for Diagnostic Differentiation in Uteri's Cervix Biopsies with Cervical/Vaginal Atypical Cytology.* Journal of Fluorescence, 2008. **18**(5): p. 979-985.
147. Chidananda, S.M., K. Satyamoorthy, L. Rai, A.P. Manjunath, and V.B. Kartha, *Optical diagnosis of cervical cancer by fluorescence spectroscopy technique.* International Journal of Cancer, 2006. **119**(1): p. 139-145.
148. Georgakoudi, I., E.E. Sheets, M.G. Müller, V. Backman, C.P. Crum, K. Badizadegan, R.R. Dasari, and M.S. Feld, *Trimodal spectroscopy for the detection and characterization of cervical precancers in vivo.* American Journal of Obstetrics and Gynecology, 2002. **186**(3): p. 374-382.
149. Chang, S.K., 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, 2005. **10**(2): p. 024031-11.
150. Mirabal, Y.N., S.K. Chang, E.N. Atkinson, A. Malpica, M. Follen, and R. Richards-Kortum, *Reflectance spectroscopy for in vivo detection of cervical*

precancer. Journal of Biomedical Optics, 2002. **7**(4): p. 587-594.
151. Marín, N.M., A. Milbourne, H. Rhodes, T. Ehlen, D. Miller, L. Benedet, R. Richards-Kortum, and M. Follen, *Diffuse reflectance patterns in cervical spectroscopy.* Gynecologic Oncology, 2005. **99**(3, Supplement 1): p. S116-S120.
152. Arifler, D., C. MacAulay, M. Follen, and R. Richards-Kortum, *Spatially resolved reflectance spectroscopy for diagnosis of cervical precancer: Monte Carlo modeling and comparison to clinical measurements.* Journal of Biomedical Optics, 2006. **11**(6): p. 064027-16.
153. Weber, C.R., R.A. Schwarz, E.N. Atkinson, D.D. Cox, C. MacAulay, M. Follen, and R. Richards-Kortum, *Model-based analysis of reflectance and fluorescence spectra for in vivo detection of cervical dysplasia and cancer.* Journal of Biomedical Optics, 2008. **13**(6): p. 064016-10.
154. Chang, V.T.-C., P.S. Cartwright, S.M. Bean, G.M. Palmer, R.C. Bentley, and N. Ramanujam, *Quantitative physiology of the precancerous cervix in vivo through optical spectroscopy.* Neoplasia, 2009. **11**(4): p. 325-332.
155. Wong, P.T., R.K. Wong, T.A. Caputo, T.A. Godwin, and B. Rigas, *Infrared spectroscopy of exfoliated human cervical cells: evidence of extensive structural changes during carcinogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(24): p. 10988-10992.
156. Wong, P.T.T., R.K. Wong, and M.F.K. Fung, *Pressure-Tuning FT-IR Study of Human Cervical Tissues.* Applied Spectroscopy, 1993. **47**(7): p. 1058-1063.
157. Wood, B.R., M.A. Quinn, F.R. Burden, and D. McNaughton, *An investigation into FTIR spectroscopy as a biodiagnostic tool for cervical cancer.* Biospectroscopy, 1996. **2**(3): p. 143-153.
158. Cohenford, M.A., T.A. Godwin, F. Cahn, P. Bhandare, T.A. Caputo, and B. Rigas, *Infrared Spectroscopy of Normal and Abnormal Cervical Smears: Evaluation by Principal Component Analysis.* Gynecologic Oncology, 1997. **66**(1): p. 59-65.
159. Fung, M.F.K., M. Senterman, P. Eid, W. Faught, N.Z. Mikhael, and P.T.T. Wong, *Comparison of Fourier-Transform Infrared Spectroscopic Screening of Exfoliated Cervical Cells with Standard Papanicolaou Screening.* Gynecologic Oncology, 1997. **66**(1): p. 10-15.
160. Chiriboga, L., P. Xie, W. Zhang, and M. Diem, *Infrared spectroscopy of human tissue. III. Spectral differences between squamous and columnar tissue and cells from the human cervix.* Biospectroscopy, 1997. **3**(4): p. 253-257.
161. Chiriboga, L., P. Xie, H. Yee, V. Vigorita, D. Zarou, D. Zakim, and M. Diem, *Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelial cells in the human cervix.* Biospectroscopy, 1998. **4**(1): p. 47-53.
162. Chiriboga, L., P. Xie, V. Vigorita, D. Zarou, D. Zakim, and M. Diem, *Infrared spectroscopy of human tissue. II. A comparative study of spectra of biopsies of cervical squamous epithelium and of exfoliated cervical cells.* Biospectroscopy, 1998. **4**(1): p. 55-59.
163. Wood, B.R., M.A. Quinn, B. Tait, M. Ashdown, T. Hislop, M. Romeo, and D. McNaughton, *FTIR microspectroscopic study of cell types and potential confounding variables in screening for cervical malignancies.* Biospectroscopy, 1998. **4**(2): p. 75-91.
164. Cohenford, M.A. and B. Rigas, *Cytologically normal cells from neoplastic cervical samples display extensive structural abnormalities on IR spectroscopy:*

Implications for tumor biology. Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(26): p. 15327-15332.
165. Romeo, M.J., M.A. Quinn, F.R. Burden, and D. McNaughton, *Influence of benign cellular changes in diagnosis of cervical cancer using IR microspectroscopy.* Biopolymers, 2002. **67**(4-5): p. 362-366.
166. Romeo, M.J., B.R. Wood, and D. McNaughton, *Observing the cyclical changes in cervical epithelium using infrared microspectroscopy.* Vibrational Spectroscopy, 2002. **28**(1): p. 167-175.
167. Romeo, M.J., B.R. Wood, M.A. Quinn, and D. McNaughton, *Removal of blood components from cervical smears: Implications for cancer diagnosis using FTIR spectroscopy.* Biopolymers, 2003. **72**(1): p. 69-76.
168. Chang, J.-I., Y.-B. Huang, P.-C. Wu, C.-C. Chen, S.-C. Huang, and Y.-H. Tsai, *Characterization of human cervical precancerous tissue through the fourier transform infrared microscopy with mapping method.* Gynecologic Oncology, 2003. **91**(3): p. 577-583.
169. Bambery, K.R., B.R. Wood, M.A. Quinn, and D. McNaughton, *Fourier Transform Infrared Imaging and Unsupervised Hierarchical Clustering Applied to Cervical Biopsies.* Australian Journal of Chemistry, 2004. **57**(12): p. 1139-1143.
170. Wood, B.R., L. Chiriboga, H. Yee, M.A. Quinn, D. McNaughton, and M. Diem, *Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone, and dysplastic squamous epithelium.* Gynecologic Oncology, 2004. **93**(1): p. 59-68.
171. Steller, W., J. Einenkel, L.-C. Horn, U.-D. Braumann, H. Binder, R. Salzer, and C. Krafft, *Delimitation of squamous cell cervical carcinoma using infrared microspectroscopic imaging.* Analytical and Bioanalytical Chemistry, 2006. **384**(1): p. 145-154.
172. Mark, S., R.K. Sahu, K. Kantarovich, A. Podshyvalov, H. Guterman, J. Goldstein, R. Jagannathan, S. Argov, and S. Mordechai, *Fourier transform infrared microspectroscopy as a quantitative diagnostic tool for assignment of premalignancy grading in cervical neoplasia.* Journal of Biomedical Optics, 2004. **9**(3): p. 558-567.
173. Podshyvalov, A., R.K. Sahu, S. Mark, K. Kantarovich, H. Guterman, J. Goldstein, R. Jagannathan, S. Argov, and S. Mordechai, *Distinction of cervical cancer biopsies by use of infrared microspectroscopy and probabilistic neural networks.* Applied Optics, 2005. **44**(18): p. 3725-3734.
174. Santos, L.F., R. Wolthuis, K. S, R.M. Almeida, and G.J. Puppels, *Fiber-optic probes for in vivo Raman spectroscopy in the high-wavenumber region.* Analytical Chemistry, 2005. **77**(20): p. 6747-6752.
175. Huang, Z., H. Lui, D.I. McLean, M. Korbelik, and H. Zeng, *Raman spectroscopy in combination with background near-infrared autofluorescence enhances the in vivo assessment of malignant tissues.* Photochemistry and Photobiology, 2005. **81**(5): p. 1219-1226.
176. Georgakoudi, I., B.C. Jacobson, J. Van Dam, V. Backman, M.B. Wallace, M.G. Müller, Q. Zhang, K. Badizadegan, D. Sun, G.A. Thomas, L.T. Perelman, and M.S. Feld, *Fluorescence, Reflectance, and Light-Scattering Spectroscopy for Evaluating Dysplasia in Patients With Barrett's Esophagus.* Gastroenterology, 2001. **120**(7): p. 1620-1629.
177. American National Standards Institute. and Laser Institute of America., *American national standard for safe use of lasers in health care facilities*. 2005,

Orlando, FL: Laser Institute of America. x, 82 p.
178. Savitzky, A. and M.J.E. Golay, *Smoothing and differentiation of data by simplified least squares procedures.* Analytical Chemistry, 1964. **36**(8): p. 1627-1639.
179. Lachenbruch, P.A. and M.R. Mickey, *Estimation of error rates in discriminant analysis.* Technometrics, 1968. **10**(1): p. 1-11.
180. Devore, J.L., *Probability and statistics for engineering and the sciences*. 7th ed. 2009, Belmont, CA: Thomson/Brooks/Cole. xvi, 735 p.
181. Deinum, G., D. Rodriguez, T.J. Romer, M. Fitzmaurice, J.R. Kramer, and M.S. Feld, *Histological Classification of Raman Spectra of Human Coronary Artery Atherosclerosis Using Principal Component Analysis.* Applied Spectroscopy, 1999. **53**(8): p. 938-942.
182. Mossman, D., *Three-way ROCs.* Medical Decision Making, 1999. **19**(1): p. 78-89.
183. Walsh, M.J., M.J. German, M. Singh, H.M. Pollock, A. Hammiche, M. Kyrgiou, H.F. Stringfellow, E. Paraskevaidis, P.L. Martin-Hirsch, and F.L. Martin, *IR microspectroscopy: potential applications in cervical cancer screening.* Cancer Letters, 2007. **246**(1-2): p. 1-11.
184. Drezek, R., M. Guillaud, T. Collier, I. Boiko, A. Malpica, C. Macaulay, M. Follen, and R. Richards-Kortum, *Light scattering from cervical cells throughout neoplastic progression: influence of nuclear morphology, DNA content, and chromatin texture.* Journal of Biomedical Optics, 2003. **8**(1): p. 7-16.
185. Engelbrecht, A.M., L. Louw, and F. Cloete, *Comparison of the fatty acid compositions in intraepithelial and infiltrating lesions of the cervix: part II, free fatty acid profiles.* Prostaglandins, Leukotrienes and Essential Fatty Acids, 1998. **59**(4): p. 253-257.
186. Preetha, A., N. Huilgol, and R. Banerjee, *Interfacial properties as biophysical markers of cervical cancer.* Biomedecine & Pharmacotherapy, 2005. **59**(9): p. 491-497.
187. Apgar, B.S., G.L. Brotzman, and M. Spitzer, *Colposcopy, principles and practice : an integrated textbook and atlas*. 2nd ed. 2008, Philadelphia: Saunders/Elsevier. xiii, 540 p.
188. Mitchell, M.F., D. Schottenfeld, G. Tortolero-Luna, S.B. Cantor, and R. Richards-Kortum, *Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis.* Obstet Gynecol, 1998. **91**(4): p. 626-631.
189. Stone, N., C. Kendall, N. Shepherd, P. Crow, and H. Barr, *Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers.* Journal of Raman Spectroscopy, 2002. **33**(7): p. 564-573.
190. Mo, J., W. Zheng, J.J.H. Low, J. Ng, A. Ilancheran, and Z. Huang, *High wavenumber Raman spectroscopy for in vivo detection of cervical dysplasia.* Analytical Chemistry, 2009. **81**(21): p. 8908-8915.
191. Jaillon, F., W. Zheng, and Z. Huang, *Half-ball lens couples a beveled fiber probe for depth-resolved spectroscopy: Monte Carlo simulations.* Applied Optics, 2008. **47**(17): p. 3152-3157.
192. Abdul-Karim, F., Y.S. Fu, J.W. Reagan, and W.B. Wentz, *Morphometric study of intraepithelial neoplasia of the uterine cervix.* Obstetrics and Gynecology, 1982. **60**(2): p. 210-214.
193. Dellas, A., H. Moch, E. Schultheiss, G. Feichter, A.C. Almendral, F. Gudat, and J. Torhorst, *Angiogenesis in Cervical Neoplasia: Microvessel Quantitation in*

Precancerous Lesions and Invasive Carcinomas with Clinicopathological Correlations. Gynecologic Oncology, 1997. **67**(1): p. 27-33.
194. Wang, L.-B., J.-G. Shen, S.-Z. Zhang, K.-F. Ding, and S. Zheng, *Amino acid uptake in arterio-venous serum of normal and cancerous colon tissues.* World Journal of Gastroenterology, 2004. **10** (9): p. 1297-1300.
195. Naber, S.P., *Molecular Pathology -- Detection of Neoplasia.* New England Journal of Medicine, 1994. **331**(22): p. 1508-1510.
196. Buschman, H.P., J.T. Motz, G. Deinum, T.J. Römer, M. Fitzmaurice, J.R. Kramer, A. van der Laarse, A.V. Bruschke, and M.S. Feld, *Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy.* Cardiovascular Pathology, 2001. **10**(2): p. 59-68.
197. Lewis, E.N., V.F. Kalasinsky, and I.W. Levin, *Near-infrared Fourier-transform Raman spectroscopy using fiber-optic assemblies.* Analytical Chemistry, 1988. **60**(24): p. 2658-2661.
198. Nie, S., K.L. Bergbauer, J.F.R. Kuck, and N.-T. Yu, *Near-infrared Fourier transform Raman spectroscopy in human lens research.* Experimental Eye Research, 1990. **51**(5): p. 619-623.
199. Williams, A.C., H.G.M. Edwards, and B.W. Barry, *Fourier transform Raman spectroscopy a novel application for examining human stratum corneum.* International Journal of Pharmaceutics, 1992. **81**(2-3): p. R11-R14.
200. Baraga, J.J., M.S. Feld, and R.P. Rava, *In situ optical histochemistry of human artery using near infrared Fourier transform Raman spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(8): p. 3473-3477.
201. Bitar, R.A., 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, 2006. **11**(5): p. 054001-8.
202. Levin, I.W. and E.N. Lewis, *Fourier transform Raman spectroscopy of biological materials.* Analytical Chemistry, 1990. **62**(21): p. 1101A-1111A.
203. Shim, M.G., *Medical Raman spectroscopy: in vivo and ex vivo tissue analysis for cancer diagnosis*, in *Medical Biophysics*. 2001, The University of Toronto: Toronto. p. 170.
204. Kendall, C., N. Stone, N. Shepherd, K. Geboes, B. Warren, R. Bennett, and H. Barr, *Identification of neoplasia with Raman spectroscopy.* The Journal of Pathology, 2003. **200**: p. 602-609.
205. Huang, Z., H. Lui, X.K. Chen, A. Alajlan, D.I. McLean, and H. Zeng, *Raman spectroscopy of in vivo cutaneous melanin.* Journal of Biomedical Optics, 2004. **9**(6): p. 1198-1205.
206. Huang, Z., E. Widjaja, W. Zheng, J. Mo, and C. Sheppard. *Classification of Raman Spectra of Colonic Tissues using Pattern Recognition Technique*. in *Biomedical Optics*. 2006: Optical Society of America.
207. Huang, Z., W. Zheng, E. Widjaja, J. Mo, and C. Sheppard. *Classification of colonic tissues using Raman spectroscopy and multivariate techniques*. 2006: SPIE.
208. Battey, D.E., J.B. Slater, R. Wludyka, H. Owen, D.M. Pallister, and M.D. Morris, *Axial Transmissive f/1.8 Imaging Raman Spectrograph with Volume-Phase Holographic Filter and Grating.* Applied Spectroscopy, 1993. **47**(11): p. 1913-1919.
209. Boyle, W.S. and G.E. Smith, *Charge coupled semiconductor devices.* Bell

System Tech. J., 1970. **49**: p. 587-593.
210. Hamamatsu, *CCD imaging area sensor*, in *S7030/S7031 Series*.
211. Shim, M.G., B.C. Wilson, E. Marple, and M. Wach, *Study of fiber-optic probes for in vivo medical Raman spectroscopy.* Applied Spectroscopy, 1999. **53**(6): p. 619-627.
212. Utzinger, U. and R.R. Richards-Kortum, *Fiber optic probes for biomedical optical spectroscopy.* Journal of Biomedical Optics, 2003. **8**(1): p. 121-147.
213. Motz, J.T., M. Hunter, L.H. Galindo, J.A. Gardecki, J.R. Kramer, R.R. Dasari, and M.S. Feld, *Optical Fiber Probe for Biomedical Raman Spectroscopy.* Applied Optics, 2004. **43**(3): p. 542-554.
214. Komachi, Y., H. Sato, K. Aizawa, and H. Tashiro, *Micro-optical fiber probe for use in an intravascular Raman endoscope.* Applied Optics, 2005. **44**(22): p. 4722-4732.
215. Komachi, Y., H. Sato, Y. Matsuura, M. Miyagi, and H. Tashiro, *Raman probe using a single hollow waveguide.* Optics Letters, 2005. **30**(21): p. 2942-2944.
216. Matsuura, Y., S. Kino, E. Yokoyama, T. Katagiri, H. Sato, and H. Tashiro, *Flexible Fiber-Optics Probes for Raman and FT-IR Remote Spectroscopy.* Selected Topics in Quantum Electronics, IEEE Journal of, 2007. **13**(6): p. 1704-1708.
217. McCreery, R.L., M. Fleischmann, and P. Hendra, *Fiber optic probe for remote Raman spectrometry.* Analytical Chemistry, 1983. **55**(1): p. 146-148.
218. Schwab, S.D. and R.L. McCreery, *Versatile, efficient Raman sampling with fiber optics*. Analytical Chemistry, 1984. **56**(12): p. 2199-2204.
219. Kong Chong, C., C. Shen, Y. Fong, J. Zhu, F.-X. Yan, S. Brush, C.K. Mann, and T.J. Vickers, *Raman spectroscopy with a fiber-optic probe.* Vibrational Spectroscopy, 1992. **3**(1): p. 35-45.
220. Myrick, M.L. and S.M. Angel, *Elimination of Background in Fiber-Optic Raman Measurements.* Applied Spectroscopy, 1990. **44**(4): p. 565-570.
221. Tanaka, K., M.T.T. Pacheco, J.F. Brennan Iii, I. Itzkan, A.J. Berger, R.R. Dasari, and M.S. Feld, *Compound parabolic concentrator probe for efficient light collection in spectroscopy of biological tissue.* Applied Optics, 1996. **35**(4): p. 758-763.
222. Owen, H., D.E. Battey, M.J. Pelletier, and J.B. Slater. *New spectroscopic instrument based on volume holographic optical elements*. in *Practical Holography IX*. 1995. San Jose, CA, USA: SPIE.
223. Lau, D.P., Z.W. Huang, H. Lui, C.S. Man, K. Berean, M.D. Morrison, and H.S. Zeng. *Raman spectroscopy for optical diagnosis in normal and cancerous tissue of the nasopharynx - Preliminary findings*. in *8th World Congress of Potodynamic Medicine*. 2001. Vancouver, Canada: Wiley-Liss.
224. Shih, W.C., K.L. Bechtel, and M.S. Feld, *Intrinsic Raman spectroscopy for quantitative biological spectroscopy part I: theory and simulations.* Optics Express, 2008. **16**(17): p. 12726-12736.
225. Pfefer, T.J., K.T. Schomacker, M.N. Ediger, and N.S. Nishioka, *Light propagation in tissue during fluorescence spectroscopy with single-fiber probes.* Selected Topics in Quantum Electronics, IEEE Journal of, 2001. **7**(6): p. 1004-1012.
226. Pfefer, T.J., K.T. Schomacker, M.N. Ediger, and N.S. Nishioka, *Multiple-fiber probe design for fluorescence spectroscopy in tissue.* Applied Optics, 2002. **41**(22): p. 4712-4721.
227. Nieman, L., A. Myakov, J. Aaron, and K. Sokolov, *Optical sectioning using a*

fiber probe with an angled illumination-collection geometry: evaluation in engineered tissue phantoms. Applied Optics, 2004. **43**(6): p. 1308-1319.
228. Schwarz, R.A., D. Arifler, S.K. Chang, I. Pavlova, I.A. Hussain, V. Mack, B. Knight, R. Richards-Kortum, and A.M. Gillenwater, *Ball lens coupled fiber-optic probe for depth-resolved spectroscopy of epithelial tissue.* Optics Letters, 2005. **30**(10): p. 1159-1161.
229. Jaillon, F., W. Zheng, and Z. Huang, *Beveled fiber-optic probe couples a ball lens for improving depth-resolved fluorescence measurements of layered tissue: Monte Carlo simulations.* Physics in Medicine and Biology, 2008. **53**(4): p. 937-951.
230. Thomsen, S., *Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions.* Photochemistry and Photobiology, 1991. **53**(6): p. 825-835.
231. Wang, L., S.L. Jacques, and L. Zheng, *MCML--Monte Carlo modeling of light transport in multi-layered tissues*. Computer Methods and Programs in Biomedicine, 1995. **47**(2): p. 131-146.
232. Welch, A.J., C. Gardner, R. Richards-Kortum, E. Chan, G. Criswell, J. Pfefer, and S. Warren, *Propagation of fluorescent light.* Lasers in Surgery and Medicine, 1997. **21**(2): p. 166-178.
233. Palmer, G.M. and N. Ramanujam, *Monte-Carlo-based model for the extraction of intrinsic fluorescence from turbid media.* Journal of Biomedical Optics, 2008. **13**(2): p. 024017-9.
234. Liu, Q., C. Zhu, and N. Ramanujam, *Experimental validation of Monte Carlo modeling of fluorescence in tissues in the UV-visible spectrum*. Journal of Biomedical Optics, 2003. **8**(2): p. 223-236.
235. Bechtel, K.L., W.-C. Shih, and M.S. Feld, *Intrinsic Raman spectroscopy for quantitative biological spectroscopyPart II: Experimental applications.* Optics Express, 2008. **16**(17): p. 12737-12745.
236. Wilson, R.H., K.A. Dooley, M.D. Morris, and M.-A. Mycek. *Monte Carlo modeling of photon transport in buried bone tissue layer for quantitative Raman spectroscopy*. in *Optics in Bone Biology and Diagnostics*. 2009. San Jose, CA, USA: SPIE.
237. Henyey, L.G. and J.L. Greenstein, *Diffuse radiation in galaxy.* Astrophys. J, 1941. **93**: p. 70-83.
238. Riedl, M.J., *Optical design fundamentals for infrared systems*. 2nd ed. Tutorial texts in optical engineering. 2001, Bellingham, Wash.: SPIE Press. xvii, 182 p.
239. Born, M. and E. Wolf, *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light*. 7th expanded ed. 1999, Cambridge ; New York: Cambridge University Press. xxxiii, 952 p.
240. Chang, S.K., D. Arifler, R. Drezek, M. Follen, and R. Richards-Kortum, *Analytical model to describe fluorescence spectra of normal and preneoplastic epithelial tissue: comparison with Monte Carlo simulations and clinical measurements.* Journal of Biomedical Optics, 2004. **9**(3): p. 511-522.
241. Eikje, N.S., Y. Ozaki, K. Aizawa, and S. Arase, *Fiber optic near-infrared Raman spectroscopy for clinical noninvasive determination of water content in diseased skin and assessment of cutaneous edema.* Journal of Biomedical Optics, 2005. **10**(1): p. 014013.
242. Nazemi, J.H. and J.F. Brennan III, *Lipid concentrations in human coronary artery determined with high wavenumber Raman shifted light.* Journal of Biomedical Optics, 2009. **14**(3): p. 034009.

243. Krafft, C., D. Codrich, G. Pelizzo, and V. Sergo, *Raman and FTIR imaging of lung tissue: bronchopulmonary sequestration.* Journal of Raman Spectroscopy, 2009. **40**(6): p. 595-603.
244. Hornung, R., T.H. Pham, K.A. Keefe, M.W. Berns, Y. Tadir, and B.J. Tromberg, *Quantitative near-infrared spectroscopy of cervical dysplasia in vivo.* Human Reproduction, 1999. **14**(11): p. 2908-2916.
245. Kondepati, V., H. Heise, and J. Backhaus, *Recent applications of near-infrared spectroscopy in cancer diagnosis and therapy.* Analytical and Bioanalytical Chemistry, 2008. **390**(1): p. 125-139.
246. McIntosh, L.M., R. Summers, M. Jackson, H.H. Mantsch, J.R. Mansfield, M. Howlett, A.N. Crowson, and J.W.P. Toole, *Towards Non-Invasive Screening of Skin Lesions by Near-Infrared Spectroscopy.* Journal of Investigative Dermatology 2001. **116**(1): p. 175-181.
247. Verkman, A., M. Hara-Chikuma, and M. Papadopoulos, *Aquaporins—new players in cancer biology.* Journal of Molecular Medicine, 2008. **86**(5): p. 523-529.
248. Wolthuis, R., M. van Aken, K. Fountas, Robinson, H.A. Bruining, and G.J. Puppels, *Determination of water concentration in brain tissue by Raman spectroscopy.* Analytical Chemistry, 2001. **73**(16): p. 3915-3920.
249. Hung, J., S. Lam, J.C. Leriche, and B. Palcic, *Autofluorescence of normal and malignant bronchial tissue.* Lasers in Surgery and Medicine, 1991. **11**(2): p. 99-105.
250. Huang, Z., W. Zheng, S. Xie, R. Chen, H. Zeng, D.I. McLean, and H. Lui, *Laser-induced autofluorescence microscopy of normal and tumor human colonic tissue.* International Journal of Oncology, 2004. **24**: p. 59-64.
251. Bensalah, K., D. Peswani, A. Tuncel, J.D. Raman, I. Zeltser, H. Liu, and J. Cadeddu, *Optical reflectance spectroscopy to differentiate benign from malignant renal tumors at surgery.* Urology, 2009. **73**(1): p. 178-181.
252. Zhang, G., S.G. Demos, and R.R. Alfano, *Far-red and NIR spectral wing emission from tissues under 532 and 632 nm photo-excitation.* Lasers in the life sciences, 1999. **9**: p. 432-440.
253. Zheng, W., W. Lau, C. Cheng, K.C. Soo, and M. Olivo, *Optimal excitation-emission wavelengths for autofluorescence diagnosis of bladder tumors.* International Journal of Cancer, 2003. **104**(4): p. 477-481.
254. Smith-McCune, K.K. and N. Weidner, *Demonstration and characterization of the angiogenic properties of cervical dysplasia.* Cancer Research, 1994. **54**(3): p. 800-804.
255. R. Richards-Kortum, E.S.-M., *Quantitative optical spectroscopy for tissue diagnosis.* Annual Review of Physical Chemistry, 1996. **47**: p. 555-606
256. Subhash, N., J.R. Mallia, S.S. Thomas, A. Mathews, P. Sebastian, and J. Madhavan, *Oral cancer detection using diffuse reflectance spectral ratio R540/R575 of oxygenated hemoglobin bands.* Journal of Biomedical Optics, 2006. **11**(1): p. 014018.
257. Brookner, C., U. Utzinger, M. Follen, R. Richards-Kortum, D. Cox, and E.N. Atkinson, *Effects of biographical variables on cervical fluorescence emission spectra.* Journal of Biomedical Optics, 2003. **8**(3): p. 479-483.
258. Richards-Kortum, R. and E. Sevick-Muraca, *Quantitative optical spectroscopy for tissue diagnosis.* Annual Review of Physical Chemistry, 1996. **47**(1): p. 555-606.
259. Majumder, S.K., M.D. Keller, F.I. Boulos, M.C. Kelley, and A.

Mahadevan-Jansen, *Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination.* Journal of Biomedical Optics, 2008. **13**(5): p. 054009-11.
260. Veld, D.C.G.d., M. Skurichina, M.J.H. Witjes, R.P.W. Duin, H.J.C.M. Sterenborg, and J.L.N. Roodenburg, *Autofluorescence and diffuse reflectance spectroscopy for oral oncology.* Lasers in Surgery and Medicine, 2005. **36**(5): p. 356-364.
261. Volynskaya, Z., A.S. Haka, K.L. Bechtel, M. Fitzmaurice, R. Shenk, N. Wang, J. Nazemi, R.R. Dasari, and M.S. Feld, *Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy.* Journal of Biomedical Optics, 2008. **13**(2): p. 024012-9.
262. Zhang, Q., M.G. Müller, J. Wu, and M.S. Feld, *Turbidity-free fluorescence spectroscopy of biological tissue.* Optics Letters, 2000. **25**(19): p. 1451-1453.
263. Yang, Y., Y. Ye, F. Li, Y. Li, and P. Ma, *Characteristic autofluorescence for cancer diagnosis and its origin.* Lasers in Surgery and Medicine, 1987. **7**(6): p. 528-532.
264. Profio, A., *Laser excited fluorescence of hematoporphyrin derivative for diagnosis of cancer.* Quantum Electronics, IEEE Journal of, 1984. **20**(12): p. 1502-1507.
265. Panjehpour, M., B.F. Overholt, J.L. Schmidhammer, C. Farris, P.F. Buckley, and T. Vo-Dinh, *Spectroscopic diagnosis of esophageal cancer: New classification model, improved measurement system.* Gastrointestinal Endoscopy, 1995. **41**(6): p. 577-581.
266. Zonios, G., L.T. Perelman, V. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, and M.S. Feld, *Diffuse Reflectance Spectroscopy of Human Adenomatous Colon Polyps In Vivo.* Applied Optics, 1999. **38**(31): p. 6628-6637.
267. Harris, A., M. Garg, X. Yang, S. Fisher, J. Kirkham, D.A. Smith, D. Martin-Hirsch, and A. High, *Raman spectroscopy and advanced mathematical modelling in the discrimination of human thyroid cell lines.* Head & Neck Oncology, 2009. **1**(1): p. 38.
268. Santos, R.S., H.A. Sidaoui, L. Silveira, C.A.G. Pasqualucci, and M.T.T. Pacheco, *Classification system of Raman spectra using cluster analysis to diagnose coronary artery lesions.* Instrumentation Science and Technology, 2009. **37**: p. 327-344.
269. Companjen, A.R., L.I. van der Wel, L. Wei, J.D. Laman, and E.P. Prens, *A modified ex vivo skin organ culture system for functional studies.* Archives of Dermatological Research, 2001. **293**(4): p. 184-190.
270. Ikuta, S., N. Sekino, T. Hara, Y. Saito, and K. Chida, *Mouse Epidermal Keratinocytes in Three-Dimensional Organotypic Coculture with Dermal Fibroblasts Form a Stratified Sheet Resembling Skin.* Bioscience, Biotechnology, and Biochemistry, 2006. **70**(11): p. 2669-2675.
271. Livera, G., G. Delbes, C. Pairault, V. Rouiller-Fabre, and R. Habert, *Organotypic culture, a powerful model for studying rat and mouse fetal testis development.* Cell and Tissue Research, 2006. **324**(3): p. 507-521.
272. Guruswamy, S., S. Lightfoot, M.A. Gold, R. Hassan, K.D. Berlin, R.T. Ivey, and D.M. Benbrook, *Effects of Retinoids on Cancerous Phenotype and Apoptosis in Organotypic Cultures of Ovarian Carcinoma.* Journal of the National Cancer Institute 2001. **93**(7): p. 516-525.
273. Ootani, A., S. Toda, K. Fujimoto, and H. Sugihara, *An Air-Liquid Interface Promotes the Differentiation of Gastric Surface Mucous Cells (GSM06) in*

Culture. Biochemical and Biophysical Research Communications, 2000. **271**(3): p. 741-746.
274. Rader, J., T. Golub, J. Hudson, D. Patel, M. Bedell, and L. Laimins, *In vitro differentiation of epithelial cells from cervical neoplasias resembles in vivo lesions.* Oncogene, 1990. **5**: p. 571.
275. Meyers, C., *Organotypic (raft) epithelial tissue culture system for the differentiation-dependent replication of papillomavirus.* Methods in Cell Science, 1996. **18**(3): p. 201-210.