Elucidation of Raman Spectral Markers in the Progression of Cervical Cancers

By

Rubina S. Shaikh

[LIFE09200904013]

Tata Memorial Centre, Mumbai

A thesis submitted to the
Board of Studies in Life Sciences

In partial fulfillment of requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Of

HOMI BHABHA NATIONAL INSTITUTE

January, 2016

HOMI BHABHA NATIONAL INSTITUTE

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Rubina Shaikh entitled **'Elucidation of Raman spectra marker in progression of cervical cancers'** and recommend that it may be accepted as fulfilling the dissertation requirements for the Degree of Doctor of Philosophy.

---------- Date: 6.1.16

Chairperson: -- Prof. Rajiv Sarin

---------- Date: 6/1/16

Guide/Convener -- Dr. C. Murali Krishna

---------- Date: 6/1/2016

Member 1: -- Dr. Shubhada Chiplunkar

---------- Date: 6/1/16

Member 2: -- Dr. Prasanna Venkatraman

---------- Date: 6/1/16

External Examiner: -- Prof. G. Krishnamoorthy

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it may be accepted as fulfilling the dissertation requirement.

Date: 6/1/16

Place: Navi Mumbai

Dr. C. Murali Krishna

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Rubina S.Shaikh

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

List of Publications arising from the thesis

Journal

1. "*In vivo* Raman spectroscopy of cervix cancers" S. Rubina, P. Sathe, T. Kumar Dora, S. Chopra, A. Maheshwari, C. Murali Krishna, *SPIE-8940, Optical Biopsy XII,* 89400E,**2014** doi:10.1117/12.2033937.
2. "*In vivo* Raman spectroscopy of Human uterine cervix: Exploration utility of vagina as an internal control" S. Rubina, T. Kumar Dora, S. Chopra, A. Maheshwari, D. Kedar K., R. Bharat, C. Murali Krishna, Journal of biomedical optics, **2014,** 19(8).
3. "Raman spectroscopic classification of exfoliated cervical cells" S. Rubina, A. Maheswari, D. Kedar, R. Bharat and C. Murali Krishna, Vibrational spectroscopy, **2013**, 68**.**
4. "Raman Spectroscopic Study on Prediction of Treatment Response in Cervical Cancers" S. Rubina, M.S.Vidyasagar and C. Murali Krishna, J Innov Opt Health Sci., **2013**, **6**.
5. "Raman spectroscopy in cervical cancers: An update" " S. Rubina and C. Murali Krishna, Journal of Cancer Research and Therapeutics, **2015**, 11(1).
6. "A comparative evaluation of diffuse reflectance and Raman Spectroscopy in the detection of cervical cancer" Vasumathi G. Prabitha, S. Rubina, S. Chopra, A. Maheshwari, K. Deodhar, B. Rekhi, N. Sukumar, C. Murali Krishna and N. Subhash, journal of biophotonics, **2016** (*in revision*)

Conference poster presentations:

1. "Raman spectroscopic study on exfoliated human cervical cells" S. Rubina, Maheswari Amita, Deodhar K. Kedar, Rekhi Bharat and C. Murali Krishna, SPEC-2012 -Shedding New Light on Disease, Chaing-Mai, Thailand.
2. "Raman Spectroscopic study on Human papillomavirus (HPV) positive and negative cell lines" S. Rubina, T. Pradhan, S.V. Chiplunkar, and C. Murali Krishna ICORS 2012, IISc Bangalore, India.
3. "Raman spectroscopic study for detection of cervical cancers" S. Rubina, Maheswari Amita , Deodhar K. Kedar , Rekhi Bharat and C. Murali Krishna, AOGIN 2011, TMH, India.
4. "Raman spectroscopic study on prediction of radio-response in cervical cancers" S. Rubina, M.S. Vidyasagar and C. Murali Krishna, AACR 2011, Gurgaon, New Delhi, India.

Other publication:

1. "Raman spectroscopic study of radioresistant oral cancer sublines established by fractionated ionizing radiation." Mohd Yasser, S. Rubina , C. Murali Krishna, Tanuja Teni, PLOS ONE, **2014.**

(Rubina S. Shaikh)

Dedicated to my Parents….

ACKNOWLEDGEMENTS

I take this opportunity to convey my gratitude to all the participants of the study. A sincere and heart filled gratitude to my mentor Dr. C. Murali Krishna for being an inspiring guide.

I am thankful to Dr. Shubhada Chiplunkar (Director, ACTREC), Dr. Sudip Gupta (Deputy Director, ACTREC) for providing an opportunity to work in this institution and the excellent infrastructure. I am fortunate to have Dr. Rajiv Sarin (Chairperson), Dr. Shubhada Chiplunkar, Dr. Prasanna Venkatraman as my doctoral committee members. Their expert comments, critical analysis of the results and helpful suggestions have contributed significantly to the work. A special thanks to Dr. Tapas, Dr. Supriya Shastri, Dr. Amita Maheshwari, Dr. Kedar Deodhar, Dr. Bharat Rekhi, for their extremely precious clinical inputs in the work. I would also like to acknowledge Dr. Supriya sonawane, Biji sister, Priyanka and Nidhul for their help during patient counseling and their co-ordination.

I want to express my warmest thanks to all past and current Chilakapati lab members Surya, Avinash, Sandeep, Mahesh, Isha, Aditi, Piyush, Priyanka, Arti and Pramod for keeping the lab environment lively and pleasant. I thankful to all my friends Renam, Devaita, Prasanna, Madhura, Snehal, Pooja, Isha, Leena, Nilofar, Gauri, and the entire student's community of ACTREC who kept the tough and tense days of my Ph.D as joyous as possible. I would like to thank all the staff from common instrument facility, photography, library, IT, administration and accounts department of ACTREC for their constant help and support. I am very much thankful to, my family (Mom, Dad, Mobin and Sana) for their constant support and encouragement. You all always have been my greatest strength and source of inspiration.

Table of Contents

Preface

A carcinoma of a uterine cervix is the fourth most common cancer among females worldwide. More than 70% of the global burden is contributed by developing countries including India. Annually, more than one million new cases are diagnosed in India with >50% mortality, which is primarily attributed to late diagnosis. The Papanicolau test (Pap test) has been used as a preliminary screening tool. An abnormal Pap smear is followed by colposcopic-guided biopsies for confirmatory diagnosis. Histopathology is the best standard for cervical cancer diagnosis. However, conventional screening and/or diagnosis tools have been known to suffer from disadvantages, such as tedious methodology, long output duration, and the inter-observer variability, besides patient discomfort. To improve the screening and/or diagnosis techniques, it is necessary to evaluate possible alternatives to present screening methodology. Raman spectroscopy can be one such alternative.

Chapter 1 gives a general introduction to the work presented in the thesis. In this chapter, the anatomy of the uterine cervix, histology along with epidemiology, etiology, and types of cervical cancers, including their staging have been discussed briefly. The current screening and diagnostic methods with their limitations have also been discussed. The literature review on application of optical spectroscopy in cervical cancer screening/diagnosis has been provided. In the later part of the chapter, emphasis is given on the biomedical application of Raman spectroscopy in diagnosis/ screening of cervical cancers; Raman effect,

instrumentation and multivariate analysis. The chapter is then concluded with the identification of aims and objectives of the study.

Even though a substantial number of *in vivo* Raman Spectroscopic studies have been carried out on cervical cancers, further validations on diverse population and in a big cohort are necessary for the translation of this technology into the clinics. Therefore, the *in vivo* Raman spectroscopic study for cervical cancer diagnosis in the Indian population was conducted. This has been presented in chapter 2. In the first section of this chapter, standardization of *in vivo* Raman setup utilizing *ex vivo* cervical tissue specimens is discussed. In the second section, the efficacy of a fiberoptic probe coupled Raman spectroscope for *in vivo* application for cervical cancer diagnosis has been explained. The last section of the chapter explains the utility of the vagina as an internal control.

In developing countries, an *in vivo* approach may not be practical, as it requires Raman instrumentation on site as well as stringent experimental conditions like dark room and regulated temperature. In these circumstances, less invasive samples like exfoliated cells may be more practical approach. Besides easy specimen collection, samples can be analyzed at a centralized facility. Thus, the Raman spectroscopic approach to differentiate normal and abnormal exfoliated cervical cell specimens and the influence of diverse factors on its classification was studied. This has been presented in the chapter 3 of the thesis. This chapter is divided into three sections: the first section describes Raman spectroscopic classification of untreated normal and abnormal specimens. The

second section of the chapter describes Raman studies of Red Blood Corpuscles (RBCs) lysis buffer treated cervical cell specimens. The second section is further divided into two parts; the first part deals with normal and abnormal specimens, whereas the second part deals with classification of treated normal, dysplastic and cancerous cell specimens. The last section of this chapter deals with the influence of lymphocytes on the classification of exfoliated cell specimens.

Human papillomavirus (HPV) is one of the major etiological factors of cervical cancer. Hence, the study to evaluate Raman spectroscopic differences in HPV positive and negative cell lines was carried out and this has been presented in chapter 4. This chapter consists of a single section in which HPV 18 positive HeLa, HPV-16 positive SiHa and HPV negative C33A cell lines were utilized in the study.

Finally the conclusions drawn from this thesis and future perspectives have been presented in chapter 5.

Synopsis

Homi Bhabha National Institute

Ph. D. PROGRAMME

1. Name of the Student:	Rubina S. Shaikh
2. Name of the Constituent Institution:	Tata Memorial Centre**,** Advanced Centre for Treatment Research and Education in Cancer
3. Enrolment No. :	LIFE09200904013
4. Title of the Thesis:	**"**Elucidation of Raman spectral marker in the progression of cervical cancers"
5. Board of study:	Life Science

Synopsis

Cancer is a major public health crisis in the world. An estimated 14 million new cancer cases and 8 million cancer deaths occurred in 2012, globally [1]. Cervical cancer is the fourth most common cancer affecting women worldwide, after breast, colorectal and lung cancers. More than 70% of the global cervical cancer burden is reported to be borne by developing countries and greater than one fifth of all new cases are diagnosed in India [1]. The prognosis of cervical cancer is determined by the stage at which the disease is presented, as determined by tumor extent, presence of lymph-node metastases and distant metastases. Cervical cancer is curable if detected early. Unfortunately, in developing countries like India, the majority of cervical cancers subjects present at advanced stages (Stage IIA and above) due to lack of stringent screening programs [2]. The Papanicolau test (Pap test), Human Papilloma Virus (HPV) testing, liquid based cytology, visual inspection of cervix after applying Lugol's iodine (VILI) or acetic acid (VIA) are well-known screening tests [2]. In routine clinical practice, an abnormal Pap smear is followed by colposcopic guided biopsies for confirmatory diagnosis. While histopathogical examination of excised biopsies remains the gold standard for cervical cancers diagnosis, current conventional screening/diagnosis tools are also known to suffer from several disadvantages like tedious sample processing, long output duration and the inter-observer variability [2].

Current research has revealed that optical screening/diagnostic methods are potential alternative/adjunct to existing cancer diagnostics. Various optical spectroscopic techniques including Raman have been explored in cancer diagnosis [3, 4].

RATIONALE AND OBJECTIVES:

Earlier studies on *in vivo* cervical cancers have demonstrated the feasibility of classifying normal and abnormal condition by Raman spectroscopy [5, 6]. Further careful validations on diverse population and larger cohorts are required for translation of this technology into clinics. **The present dissertation aims to evaluate the efficacy of Raman spectroscopic methods for non-invasive/minimal-invasive and objective screening/diagnosis of cervical cancers under clinical setting.**

The thesis focuses on the following objectives:

1. To characterize Raman spectral differences between normal and cancerous cervical tissues, in both, *in-vivo* and *ex- vivo* conditions.
2. To characterize Raman spectral differences between normal, pre-cancerous and cancerous cervical exfoliated cells.
3. To characterize Raman spectral differences in HPV expressing and non expressing cell-lines

1. Objective 1:*Ex vivo* and *in vivo* Raman spectroscopic study on cervical cancers : Previous *ex vivo* studies on cervical cancer have demonstrated the potential of Raman spectroscopic methods in classifying normal, premalignant and tumor conditions [7, 8]. This objective was taken up to evaluate the reproducibility of spectra features using fiberoptic probe coupled Raman system as well as to standardize the protocol for *in vivo* studies. This objective was carried out in two parts, the first was to confirm the reproducibility of spectra

using tissue biopsies and the second part was to implement it for *in vivo* cervical cancer studies in a clinical setting.

***1.1. Ex vivo study: confirmation of reproducibility of spectra*:** Spectra from pathologically certified, 27 cervix biopsies (tumor and normal) were acquired. From these, 16 tumor tissues were collected from locally advanced cancer subjects before undergoing treatment and normal tissues were collected from 11 subjects undergoing hysterectomy. Tissues were snap frozen in liquid nitrogen and stored at -80^{0}C until use. Spectra were acquired using a HE-785 commercial Raman system (Jobin-Yvon-Horiba, France). Briefly, this system consists of a diode laser HE-785 as excitation source (wavelength-785 nm), and a HE-785 spectrograph (HE-785, HORIBA Jobin Yvon, France) coupled with a CCD (CCD-1024X256-BIDD-SYN, Synapse) as dispersion and detection elements, respectively. Spectral acquisition parameters were: laser power at sample ~80 mW, integration-15 seconds and 3-accumulations. Pre-processing of raw spectra was carried out by a standard procedure which involves rectification for CCD response with a NIST certified standard reference material-2241 (SRM- 2241) followed by subtraction of background signals due to optical elements. Pre-processed spectra were used for Principal Components Linear Discriminating analysis (PC-LDA) using algorithms implemented in MATLAB (Mathworks Inc.) based in-house software. Standard models of normal and tumor were developed using 148 and 201 spectra from 11 normal and 16 tumor tissues, respectively. Leave one out cross validation (LOOCV) yielded sensitivity and specificity of 94 and 91%. Corroborating earlier observations; mean spectrum of normal conditions

was dominated by collagen bands while non-collagenous proteins and nucleic acid were predominant in tumor spectrum [7-9]. The overall findings of the study confirmed the reproducibility of spectral features.

1.2. <u>In vivo study: Development of standard model, cross validation and evaluation with independent test data</u>

1.2.1. Classification between normal and cancer: The above standardized spectral acquisition and data analysis protocols were used for the *in vivo* studies; *in vivo* Raman spectra from normal cervix, cancerous lesion and vaginal sites of 103 subjects were acquired. Spectral acquisition parameters were: λex-785 nm, laser power-80 mW, spectra were integrated for 5 seconds and averaged over 3 accumulations. Spectra were preprocessed as per the earlier described procedure. The mean spectra of normal cervix and vaginal sites exhibit characteristic spectral features of amide III and strong and broad amide I, which can be attributed to collagenous proteins. Prominent features of tumor, with respect to normal spectrum, are strong and sharper amide I, minor shifts in δ CH_2 and a distinct band at 1340 cm^{-1} which are indicative of DNA and non-collagenous proteins. These findings corroborate earlier *ex vivo* and *in vivo* cervical. Raman spectroscopic studies [7-9]. Pre-processed Raman spectra of tumor and normal spectra were subjected to PC-LDA. PC-LDA gave average classification efficiency of 98.5% [10].

***1.2.2. Exploring utility of vagina* as an internal control:** Since cervical cancer subjects present at advanced stages (Stage IIA and above) in developing countries like India, the majority of the cervix is diseased; no normal cervix sites are

obtainable to acquire control spectra. It has been reported that variability due to menopausal status, hormonal status, age and parity may lead to bias, thus there is a need for another internal control. It is also known that the composition of the vagina and ectocervix are similar, as they contain an inner lining of squamous epithelial cells. Thus the vagina can serve as a good internal control. Hence we have explored the utility of the vagina as an internal control. This approach could be helpful to circumvent inter-patient variability due to menopausal status, hormonal status, age, parity, and it could be especially useful in screening camps where Colposcopy may not be available.

A. Classification among controls: To explore the variations between the control groups, spectra of normal cervix, vaginal sites of normal and tumor subjects were analyzed by PC-LDA. The higher misclassification was observed between the spectra of all the control groups (i.e normal cervix, vaginal sites of normal and tumor subjects) is indicative of the biochemical similarities among these groups. As spectra of vagina of normal and tumor subjects show similarity, we have grouped them together and referred to them as vagina spectra in our subsequent evaluation of vagina as control.

B. Evaluating utility of vagina as an internal control: In order to evaluate the efficacy of Raman spectroscopic methods in discriminating tumor conditions from control groups, spectra of tumor, normal cervix, and vagina were subjected to PC-LDA. PC-LDA exhibited high classification among the clusters belonging to tumor and control spectra, whereas spectra from normal cervix and vaginal sites exhibited very high overlap. The high misclassifications between normal

cervix and vaginal sites once again suggest similarities between the biochemical compositions. Suggesting that vagina can be used as internal control. The findings of the study corroborate with earlier studies and suggest the applicability of Raman spectroscopic methods for objective, noninvasive and rapid diagnosis of cervical cancers [5, 7-10]. The study also demonstrates that Raman spectroscopy may be used for improving cervical cancer diagnosis by incorporating internal control like vagina to circumvent the influence of parameters like hormonal status, menopausal status, and age; as well as the requirement of colposcope especially for mass screening camps [10, 11].

2. Objective 2:Exploring Raman spectral features of exfoliated normal and abnormal cervical exfoliated cells : Several FTIR reports have demonstrated that normal and cancerous exfoliated cell specimens can be distinguished [12-15]. But water, a universal constituent of cell specimens, is a serious hurdle in FTIR spectroscopy [16]. Dried specimens were used to overcome this problem. It is well known that the morphology and biochemical composition of cells is altered due to drying [16]. The vibrational characteristics of such sample might not represent the true biological state of the cells. Moreover, since specimens are dried, the same cells could not be used for Pap staining. Use of a parallel sample for Pap staining may not be ideal for cytological correlation, as abnormal cell content in an ‘abnormal’ smear can vary. Raman spectroscopy requires minimal or no sample preparation. Hence, both Raman spectroscopy and Pap staining can be employed on the same specimen, which in turn can lead to better cytological correlation. So far, to the best of our knowledge, no studies toward classification

of normal and abnormal specimens using certified cell specimens by RS have been reported. The present objective aims at exploring the potential of RS in classifying normal and abnormal exfoliated cervical cells. Exfoliated cervical cell specimens from 107 subjects were collected, out of which 94 specimens having good cell yield were included in study.

2.1 Classification of normal and abnormal cervical exfoliated cell specimens: In the first approach; out of 94 specimens, spectral data from 37 cell specimens were acquired and analyzed. Raman mean spectrum of normal cell specimen showed predominant protein features indicated by bands at amide I, δCH_2 stretch and ring breathing mode of phenylalanine. In contrast to the normal specimen spectrum, the mean spectra of abnormal specimens shows strong features associated with blood components like fibrin and Red Blood Corpuscles (RBCs) indicated by the C-C stretching mode of heme, fibrin, δCH_2, C-C symmetrical stretch from heme and phenylalanine [16]. Pap stained slides also exhibited the presence of blood in abnormal specimens. Further, PC-LDA yielded classification efficiencies of 86% and 84% for normal and abnormal specimens, respectively.

2.2 Classification of RBC lysed normal and abnormal cervical exfoliated cell specimens: As the presence of RBCs in a specimen can also occur in non cervical cancerous conditions, it is not an ideal marker for classification of specimens. In order to avoid the presence of RBCs in specimens, the remaining 57 specimens were treated with RBC lysis buffer. The mean spectrum of normal and abnormal cell specimens showed predominant protein features indicated by bands at amide I, δCH_2 stretch and ring breathing mode of phenylalanine. Differences in amide I,

amide III, δCH_2 and 1000–1200 cm^{-1} regions were observed [16]. Further, PC-LDA resulted in classification efficiency of 79% and 78% for normal and abnormal smear, respectively. Misclassifications in both the approaches can be attributed to the presence of normal cells in abnormal specimens.

2.3 Classification of normal, precancerous and cancerous cervical exfoliated cell specimens: In the last approach, classification among normal, precancerous and cancerous exfoliated cell specimens was also explored. It was observed that precancerous spectra showed overlap with normal and cancerous, whereas cancerous and normal grouped showed the tendency of classification.

2.4 Influence of lymphocytes on classification of normal and abnormal cervical exfoliated cell specimens: Furthermore, the lymphocyte influence on the classification of exfoliated cell specimens was also evaluated, the other popular confounding factor. Raman spectra were recorded from exfoliated cervical cell specimens, lymphocytes and different ratios of exfoliated cells + lymphocytes. After Raman spectral acquisition cell pellets were smeared and Pap stained for conformational diagnosis. Spectra in the 900-1800 cm^{-1} region were utilized for classification using PCA. Two exclusive clusters for lymphocytes and exfoliated cell specimens were observed. Also, spectra of 1:1, 1:2 and 1:3 compound specimens showed overlap with exfoliated cell specimens, which indicates minimal or no influence of lymphocytes on classification. This was also further confirmed by the PCA of exfoliated cell specimens with different lymphocyte concentrations. The findings suggest that the presence of lymphocytes showed minimal or no influence on the classification of cell specimens.

Objective 3: <u>Exploring Raman spectral features of HPV expressing and non expressing cell lines:</u> High risk human papillonavirus (HR-HPV) is well known etiological factor of cervical cancers. Persistence of the virus is linked to the development of a high-grade precursor lesion or "precancer". Although the presence of HPV has clinical significance, it is pertinent to note that all HPV infection may not lead to cervical cancers after clearance of HPV infection [17]. Recently, high-risk HPV strains testing have been incorporated into routine cervical cancer screening for menopausal females in developed countries [1]. Although it has been reported that HPV infection related changes can be detected by Raman spectroscopy, this objective was undertook for evaluating the efficacy of our Raman spectroscopy to classify HPV positive and negative cell lines [18, 19]. Spectra of HPV 18 positive HeLa, HPV 16 positive SiHa and HPV negative C33A cell lines were acquired. HPV negative (C33A) cells and HPV positive (HeLa and SiHa) cells showed distinct differences at amide I, δCH_2 region. Minor variations in amide III region were also observed, whereas no significant differences between HPV positive cells were discernible. A possible explanation for this observation could be that HPV infection eventually leads to oncoprotein expression resulting in differences in protein compositions in the host cells. PC-LDA gave well separated clusters with classification efficiency of ~95%. The findings of the study corroborate earlier reports and demonstrate subtle but significant differences between HPV positive and HPV negative cell lines, which can be differentiated using Raman spectroscopy [19].

References

1. Ferlay J.S.I., M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, F. Bray, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer. GLOBOCAN 2012 v10. 2013.
2. Kerkar R.A. and Y.V. Kulkarni, Screening for Cervical Cancer: An Overview. J Obstet Gynecol India, 2006. **56**(2):115-122.
3. Evers D., B. Hendriks, G. Lucassen and T. Ruers, Optical Spectroscopy: Current Advances and Future Applications in Cancer Diagnostics and Therapy. Future Oncol, 2012. **8**(3):307-320.
4. Ellis D.I., D.P. Cowcher, L. Ashton, S. O'Hagan and R. Goodacre, Illuminating Disease and Enlightening Biomedicine: Raman Spectroscopy as a Diagnostic Tool. Analyst, 2013. **138**(14):3871-3884.
5. 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):427-431
6. Duraipandian S., W. Zheng, J. Ng, J.J. Low, A. Ilancheran and Z. Huang, Near-Infrared-Excited Confocal Raman Spectroscopy Advances in Vivo Diagnosis of Cervical Precancer. Journal of Biomedical Optics, 2013. **18**(6):067007-067007.
7. 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):123-132.
8. Krishna C.M., N. Prathima, R. Malini, B. Vadhiraja, R.A. Bhatt, D.J. Fernandes, P. Kushtagi, M. Vidyasagar and V. Kartha, Raman Spectroscopy Studies for Diagnosis of Cancers in Human Uterine Cervix. Vibrational spectroscopy, 2006. **41**(1):136-141.

9. Rubina S., M.S. Vidyasagar, C. M. Krishna, Raman Spectroscopic Study on Prediction of Treatment Response in Cervical Cancers, J Innov Opt Health Sci., 2013, **6**:1350014.
10. Rubina S., P. Sathe, T. K. Dora, S. Chopra, A. Maheshwari, C. M. Krishna, *In vivo* Raman Spectroscopy of Cervix Cancers. Proc. SPIE 8940, Optical Biopsy XII,, 2014. **8940**.
11. Rubina S., T.K. Dora, S. Chopra, A. Maheshwari, K.D. Kedar, R. Bharat and C.M. Krishna, *In Vivo* Raman Spectroscopy of Human Uterine Cervix: Exploring the Utility of Vagina as an Internal Control. J Biomed Opt, 2014. **19**(8):087001.
12. Wong P., 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, 1991. **88**(24):10988-10992.
13. 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, 1998. **95**(26):15327-15332.
14. Sindhuphak R., S. Issaravanich, V. Udomprasertgul, P. Srisookho, S. Warakamin, S. Sindhuphak, R. Boonbundarlchai and N. Dusitsin, A New Approach for the Detection of Cervical Cancer in Thai Women. Gynecol Oncol, 2003. **90**(1):10-14.
15. 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):69-76.
16. Rubina S., A. Maheshwari, K. Deodar, R. Bharat and C.M. Krishna, Raman Spectroscopic Study on Classification of Cervical Cell Specimens. Vibrational spectroscopy, 2013. **68**:6.
17. Chen H.C., M. Schiffman, C.Y. Lin, M.H. Pan, S.L. You, L.C. Chuang, C.Y. Hsieh, K.L. Liaw, A.W. Hsing and C.J. Chen, Persistence of Type-Specific

Human Papillomavirus Infection and Increased Long-Term Risk of Cervical Cancer. J Natl Cancer Inst, 2011. **103**(18):1387-1396.

18. Jess P.R., D.D. 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):2723-2728.
19. Ostrowska K.M., A. Malkin, A. Meade, J. O'Leary, C. Martin, C. Spillane, H.J. Byrne and F.M. Lyng, Investigation of the Influence of High-Risk Human Papillomavirus on the Biochemical Composition of Cervical Cancer Cells Using Vibrational Spectroscopy. Analyst, 2010. **135**(12):3087-3093.

List of publications:

a) Published:

[1] "*In vivo* Raman spectroscopy of cervix cancers" S. Rubina, Priyanka Sathe, Tapas Kumar Dora, Supriya Chopra, Amita Maheshwari, C. Murali Krishna, *SPIE-8940, Optical Biopsy XII,* 89400E (2014) doi:10.1117/12.2033937.

[2] "Raman spectroscopic classification of exfoliated cervical cells" S. Rubina, Maheswari Amita, Deodhar K. Kedar, Rekhi Bharat and C. Murali Krishna, Vibrational spectroscopy **68**, 115-121(2013).

[3] "Raman Spectroscopic Study on Prediction of Treatment Response in Cervical Cancers" S. Rubina, M.S.Vidyasagar and C. Murali Krishna, J Innov Opt Health Sci., **6**, 1350014, (2013)

b) Communicated:

[1] "*In vivo* Raman spectroscopy of Human uterine cervix: Exploration utility of vagina as an internal control" S. Rubina, Tapas Kumar Dora, Supriya Chopra, Amita Maheshwari, Deodhar Kedar K., Rekhi Bharat, C. Murali Krishna, Journal of biomedical optics (under revision).

c) Other publication:

[1] "Raman spectroscopic study of radioresistant oral cancer sublines established by fractionated ionizing radiation." Mohd Yasser, S. Rubina , C. Murali Krishna , Tanuja Teni, PLOS ONE (2014) (in press)

Signature of Student:
Date: 3rd june 2014

Doctoral Committee:

S. No.	Name	Designation	Signature	Date
1	Dr.Rajiv Sarin	Chairperson		3/6/14
2	Dr. C. Murali Krishna	Convener		6/6/14
3	Dr. Shubhada Chiplunkar	Member		27/6/14
4	Dr. Prasanna Venkatraman	Member		5/6/2014

Dr. S.V. Chiplunkar
Director, Chairperson Academics and
Training Program, ACTREC

Prof. K. Sharma
Director, Academics
Tata Memorial Centre

Dr. S. V. Chiplunkar
Director
Advanced Centre for Treatment, Research & Education in Cancer (ACTREC)
Tata Memorial Centre
Kharghar, Navi Mumbai 410210.

Prof. K.S. Sharma
DIRECTOR - ACADEMICS, TMC
Mumbai - 400 012

Abbreviations

µm	micrometer
ACTREC	Advanced Centre for Treatment, Research and Education in Cancer
ANN	artificial neural networks
ASC-H	Atypical squamous cells- cannot exclude HSIL
ASC-US	Atypical squamous cells of undetermined significance
BMI	body mass index
CaF_2	Calcium fluoride
CCD	Charged Couple Devise
CIN-1	Cervical intraepithelial neoplasia-1
CIN-2	Cervical intraepithelial neoplasia-2
CIN-3	Cervical intraepithelial neoplasia-3
D.P.X	Dibutyl Phathalate Xylene
dCV	double cross-validation
DNA	deoxyribonucleic acid
FAD	Flavin Adenine Dinucleotide
FFPP	formalin-fixed paraffin preserved
FIGO	International Federation of Gynecology and Obstetrics
FP	fingerprint
FTIR	Fourier transform infra-red absorption
GA-PLSDA	genetic algorithm-partial least squares-discriminant analysis
HCA	hierarchical cluster analysis
HR-HPV	High risk -Human papilloma virus
HSIL	High-grade squamous intraepithelial lesion
HW	high-wavenumber
IR	Infrared
IRB	Institutional Review Board
IUD	intrauterine contraceptive device
LDA	linear discriminant analysis
LOOCV	leave-one-spectrum-out, cross-validation
LSIL	Low grade squamous intraepithelial lesion
MDC	multispectral digital colposcopy
ml	milli leter
MRDF	maximum representation and discrimination feature
MVA	Multivariate analysis
mW	Milliwatt
N	Normal cervix site
NA	Numerical aperture
NADH	Nicotinamide adenine dinucleotide dehydrogenase
NIPALS	Non-linear Iterative Partial Least Squares

Abbreviations

NIR	Near Infrared
NIST	National Institute of Standards and Technology
nm	nanometer
NOS	not otherwise specified
NVN	vaginal sites of subjects with normal cervix
NVT	vaginal sites of subjects with tumor cervix
Pap test	Papanicolaou test
PCA	Principal Component Analysis
PC-LDA	Principal Component-Linear Discriminant Analysis
PCs	principal components
PHK	primary human keratinocytes
PNN	probabilistic neural networks
RBCs	Red Blood Corpuscles
rpm	revolutions per minute
RS	Raman spectroscopy
RSS	Resonance Enchanced Raman Scattering
s	seconds
SCC	Squamous cell carcinoma
SERS	Surface Enhanced Raman Scattering
SMLR	sparse multinomial logistic regression
SR-IR	synchrotron infrared
SRM 2241	Standard Reference Material 2241
T	Tumor cervix
V	Vaginal site
VIA	Visual inspection of cervix after applying acetic acid
VILI	Visual inspection of cervix after applying Lugol's iodine

List of Figures

List of Figures

List of Tables

1 INTRODUCTION

Chapter 1

"Cervical cancers can have devastating effects with a very high human, social, and economic cost, affecting women in their prime. But this disease should not be a death sentence, even in poor countries," as very well said by Dr R. Sankaranarayanan [1].

In particular, 5,28, 000 new cervical cancer cases are detected annually, leading it to be the fourth most frequent women cancer worldwide, after breast, colorectal and lung cancers [1]. It has been reported that the 70 % of global cervical cancer burden is contributed by the developing countries and more than one fifth of all new cases are diagnosed in India. Late detection is considered to be prime cause of mortality in the developing world, due to lack of stringent screening programs [1]. Therefore, substantial efforts are needed to improve techniques to prevent this cancer. The prevention of cervical cancers usually has three checkpoints- screening, diagnosis and therapy. The current standard screening technique is Papanicolaou (Pap) test. Abnormal Pap reports are commonly followed by colposcopic examinations. Histopathology remains the gold standard for cervical cancer diagnosis. This is a relatively invasive procedure leading to the increase in diagnostic time and the cost [2].

Therefore, relatively non-invasive optical spectroscopic techniques like fluorescence, reflectance, infrared and Raman spectroscopy can be useful. They are similar to each other due to qualities like minimal/ noninvasive, real-time, objective detection of biochemical and molecular changes within a tissue. Due to these qualities, these techniques have emerged as promising techniques to aid in cancer prevention. In particular, Raman spectroscopy has been applied for the

detection of many cancers including cervical cancers [3]. In this study, the efficacy of Raman spectroscopy for *ex vivo* and *in vivo* detection of cervical cancers utilizing minimal/less invasive approaches were explored.

This chapter introduces the anatomy and histology of the cervix, epidemiological aspects, including incidence of cervical cancers, followed by an introduction to Raman effect and Raman spectroscopic instrumentation. Subsequently, a review of literature on optical spectroscopic techniques (i.e., fluorescence, reflectance, infrared spectroscopy and Raman spectroscopy) for the detection of cervical cancers is presented. Finally, the aims, objectives and structure of the thesis are presented.

1.1 Cervical cancers - An overview

1.1.1 Anatomy and histology of the cervix

The cervix is approximately 2.5 to 3 cm in length and forms the lower third of the uterus, extending into vagina. The part of the cervix that projects into the vagina is called as ectocervix, and is covered by non-keratinized stratified squamous epithelium. The uterine part of the cervix is called endocervix and is covered with mucus-secreting columnar epithelium. A pictorial depiction of the cervical anatomy is shown in Figure 1.1[4]. The cervical transformation zone is an area of metaplasic tissue between the squamous epithelium of the vagina and the glandular tissue of the endocervical canal. The squamocolumnar junction has a unique susceptibility to high risk Human papilloma virus (HPV), HPV-induced neoplastic transformation leading to cancer [5].

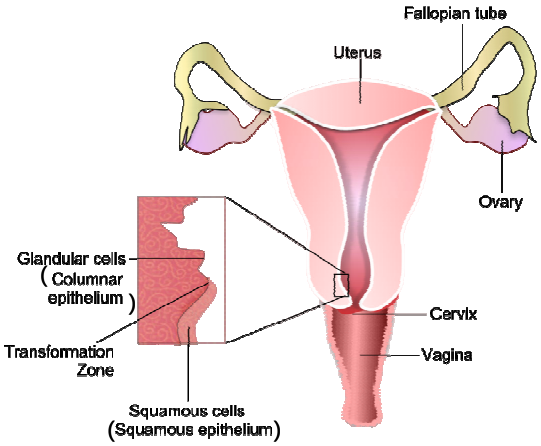


Figure 1.1: Pictorial depiction of cervical [4]

The squamous epithelium is 15 to 20 cells thick, approximately 200-400 microns in size. As basal cells mature, these cells migrate to the surface, accumulate glycogen in the cytoplasm and are compressed to acquire a flatten shape. Underlying the basal layer is the stroma, which is rich in collagen [6].

1.1.2 Incidence of cervical cancers

Cervical cancer is the fourth most common cancer affecting women worldwide, with an estimated 5,28,000 new cases annually. It is also the fourth most frequent cause of cancer death (266 000 deaths in 2012) among women worldwide [1]. More than 70% of the global cervical cancer burden is contributed by developing countries whereas more than one fifth of all new cases are diagnosed in India [1].

In sub-Saharan Africa, annually, 34.8 per 1, 00,000 women are diagnosed with cervical cancers and 22.5 per 1, 00, 000 women die due to cervical cancer, while in North America these figure are 6.6 and 2.5 per 1, 00, 000 women, respectively [1]. Lack of access to effective screening and to services that

facilitate early detection and treatment are possible reasons of these geographical differences [1].

1.1.3 Etiology of cervical cancers

A variety of etiological factors have been found to be associated with cervical carcinogenesis, such as High risk- Human Papillomavirus (HR-HPV) infection, hormonal contraceptives, smoking, parity, hygiene, co-infection with other sexually transmitted agents, genetic and immunological factors [7]. Among all the mentioned factors, HR-HPV infection has been known as the key cause of cervical cancers [5]. However, it is also known that all HR-HPV infections may not lead to cervical dysplasia. Recently, it has been reported that stages in cervical carcinogenesis include HR-HPV infection, persistence of infection, development of a high-grade precursor lesion or precancer and invasion [5].

1.1.4 Types of cervical cancers

Cervical cancers are histopathologically subdivided into four subtypes, namely squamous cell carcinoma, adenocarcinoma, mixed adenosquamous and neuroendocrine carcinoma [2]. Among these, the squamous cell carcinoma is the most common type (>90%) and it originates from the squamous epithelium of the exocervix. Adenocarcinoma arises from glandular cells of the endocervix and contributes less than 10% of tumors. Mixed adenosquamous carcinomas of cervix is the rarest among all the cervical carcinomas and aggressive subtypes [8].

1.1.5 Cervical cancer staging

Cervical cancer is staged according to the International Federation of Gynecology and Obstetrics (FIGO) staging system. This system is based on clinical examination and following are the stages according to FIGO [9].

Stage I: The carcinoma is strictly confined to the cervix.

Stage IA: Invasive cancer identified only microscopically. (All gross lesions even with superficial invasion are Stage IB cancers.) Invasion is limited to measured stromal invasion with a maximum depth of 5 mm and no wider than 7 mm.

Stage IA1: Measured invasion of stroma ≤3 mm in depth and ≤ 7 mm width.

Stage IA2: Measured invasion of stroma >3 mm and <5 mm in depth and ≤7 mm width.

Stage IB: Clinical lesions confined to the cervix greater than stage IA.

Stage IB1: Clinical lesions no greater than 4 cm in size.

Stage IB2: Clinical lesions >4 cm in size.

Stage II: Carcinoma extends beyond the uterus, but has not extended onto the pelvic wall or to the lower third of vagina.

Stage IIA: Involvement of up to the upper 2/3 of the vagina. No obvious parametrial involvement.

Stage IIA1: Clinically visible lesion ≤4 cm

Stage IIA2: Clinically visible lesion >4 cm

Stage IIB: Obvious parametrial involvement but not onto the pelvic sidewall.

Stage III: Carcinoma has extended onto the pelvic sidewall. On rectal examination, there is no cancer-free space between the tumor and pelvic sidewall. The tumor involves the lower third of the vagina. All cases of hydronephrosis or

non-functioning kidney should be included unless they are known to be due to other causes.

Stage IIIA: Involvement of the lower vagina but no extension onto pelvic sidewall.

Stage IIIB: Extension onto the pelvic sidewall, or hydronephrosis/non-functioning kidney.

Stage IV: The carcinoma has extended beyond the true pelvis or has clinically involved the mucosa of the bladder and/or rectum.

Stage IVA: Spread to adjacent pelvic organs.

Stage IVB: Spread to distant organs.

1.1.6 Current screening methods

Presently, conventional cervical screening includes Pap test/ cervical cytological examinations followed by triaged women for HPV testing in the case of atypical squamous cells of undetermined significance (ASC-US). This is further followed by identifying women for colposcopy guided biopsies with abnormal Pap results [5].

Pap test is the prime screening test for cervical cancers. This test was first described by Papanicolaou and Traut in 1943 [6]. Liquid based cytology, HPV testing, visual inspection of cervix after applying Lugol's iodine (VILI) or acetic acid (VIA) are also few other screening techniques [10, 11]. However, the conventional Pap test is the well-known method by which cervix is sampled. Exfoliated cell specimens are smeared on glass slides and fixed with fixative. Smears are further stained by nuclear and cytoplasmic stains. Smears are

categories by utilizing standard cervical cytology reporting methodology termed as Bethesda system [12].

Current interpretations of cervical smears (The Bethesda system 2001)

Specimen type

Indicate

- conventional smear (Pap smear)
- vs. liquid-based preparation
- vs. other.

Specimen adequacy

- Satisfactory for evaluation (describe presence or absence of endocervical/transformation zone component and any other quality indicators, e.g., partially obscuring blood, inflammation, etc.)
- Unsatisfactory for evaluation (specify reason)
- Specimen rejected/not processed (specify reason)
- Specimen processed and examined, but unsatisfactory for evaluation of epithelial abnormality because of (specify reason)

General categorization (optional)

- Negative for Intraepithelial Lesion or Malignancy
- Epithelial Cell Abnormality
- Other

Negative for intraepithelial lesion or malignancy

(When there is no cellular evidence of neoplasia, state this in the General categorization above and/or in the Interpretation/Result section of the report, whether or not there are organisms or other non-neoplastic findings).

(a) Organisms

- Trichomonas vaginalis
- Fungal organisms morphologically consistent with Candida species

- Shift in flora suggestive of bacterial vaginosis
- Bacteria morphologically consistent with Actinomyces species
- Cellular changes consistent with Herpes simplex virus

(b) Other non-neoplastic findings (Optional to report; list not inclusive)

- Reactive cellular changes associated with: inflammation (includes typical repair), radiation and intrauterine contraceptive device (IUD)
- Glandular cells status post hysterectomy
- Atrophy

(c) Others

- Endometrial cells (in a woman > 40 years of age) (Specify if 'negative for squamous intraepithelial lesion')

Epithelial cell abnormalities

(a) Squamous cell

- Atypical squamous cells (ASC)
- of undetermined significance (ASC-US)
- cannot exclude high-grade squamous intraepithelial lesion (ASC-H)
- Low grade squamous intraepithelial lesion (LSIL)
- encompassing: HPV/mild dysplasia/Cervical intraepithelial neoplasia (CIN) 1
- High grade squamous intraepithelial lesion (HSIL)
- encompassing: moderate and severe dysplasia, Carcinoma in situ , (CIN 2 and CIN 3)
- Squamous cell carcinoma (SCC)

(b) Glandular cell

Atypical

- endocervical cells (not otherwise specified (NOS) or specify in comments),
- endometrial cells (NOS or specify in comments),
- glandular cells (NOS or specify in comments)

Atypical

- endocervical cells, favor neoplastic
- glandular cells, favor neoplastic

Endocervical adenocarcinoma in situ

Adenocarcinoma:

- endocervical
- endometrial
- extrauterine
- not otherwise specified (NOS)

Others (List not comprehensive)

- Endometrial cells in a woman > 40 years of age

Colposcopy and the guided biopsy

Generally, abnormal Pap cases are referred for coloposcopic examinations. Colposcopy is the visual examination of the cervix using low-powered microscope, known as a colposcope. A solution of 4% acetic acid is swabbed on to the surface of cervix, causing the area of abnormal epithelium to change its color to white. To enhance the contrast at the transitional zone, an iodine solution is used to stain the normal squamous epithelium, whereas the normal columnar epithelium and abnormal epithelium do not take the stain. To confirm the diagnosis, colposcopic findings are often confirmed with one or more biopsies. Results of the histopathology serve as the gold standard for cervical precancer or cancer diagnosis. However, the need to confirm diagnosis with biopsy amplifies patient inconvenience, pain, and cost of disease management.

1.2 Optical spectroscopy in cervical cancer diagnosis

Optical spectroscopic techniques have been explored in various diseased conditions including cervical cancers due to their potentials to provide real-time, rapid, non-destructive, biochemical/molecular information.

1.2.1 Fluorescence spectroscopy

Fluorescence spectroscopy is based on the detection of luminescence of natural endogenous fluorochroms and/or exogenous fluorescencent agents. This technique is centered on particular molecule signal upon illumination of specimens with the light of specific wavelength [13]. These molecules absorb the energy and get excited from the ground state. Upon de-excitation, the molecules fluoresce to generate light with wavelength different to the excitation wavelength. The Jablonski diagram illustrating the phenomenon of fluorescence is shown in Figure 1.2. The intensity and shape of the fluorescence spectrum depends on the type and concentrations of the fluorophores in the specimen.

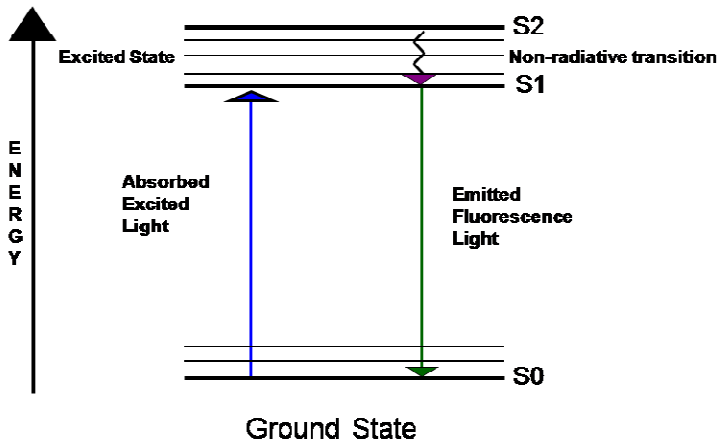


Figure 1.2: Jablonski diagram illustrating the phenomenon of fluorescence.

(http://chemwiki.ucdavis.edu/Physical_Chemistry/Spectroscopy/Electronic_Spectroscopy/Fluorescence)

Fluorescence spectroscopy is a widely investigated in several cancers including cervical cancers [14, 15]. Lohmann *et al*. reported that the intensity of Nicotinamide adenine dinucleotide dehydrogenase (NADH) band increases with the progression of tissue dysplasia and is very low in tumor tissue [16]. The findings were corroborated by the study carried out on cervix tissue sections [17]. Different groups reported the feasibility of fluorescence spectroscopy to discriminate cervical cancer tissue [18-20]. The study by Koumantakis *et al*. described various bands at 558, 583, 600, 630 and 697 nm as cervical malignancy predictors with 420 nm excitation [21]. In the following years, Ramanujam *et al*. carried out extensive work on the application of fluorescence spectroscopy for the *in vivo* detection of cervical dysplasia [15, 22-26] . Subsequent studies focused on investigation of the different variables associated with *in vivo* fluorescence spectroscopic clinical trials on cervical precancer diagnosis. The effect of variables on cervical cancer diagnosis, like tissue type, size, population, optimum excitation wavelength, and signal to noise ratio, fiber probe pressure, device, acetic acid application, cervix mucus, inter-patient variation, menstrual cycle, hormones, menopause and age were studied using fluorescence spectroscopy [27-37].

Fluorescence imaging has been explored to study the fluorescence properties associated with cervical dysplasia. In the normal cervix, high fluorescence intensity of stroma was observed as compared to the epithelium with increasing patient age [38]. Subsequent studies exhibited correlation between fluorescence image pattern with tissue premalignancy and malignancy. Several groups reported

an increase in Nicotinamide Adenine Dinucleotide (NADH) fluorescence and a reduction in collagen fluorescence with tissue malignant alterations [39-43]. In addition to these bio-molecules, increased Flavin Adenine Dinucleotide (FAD) and decreased keratin fluorescence in the epithelium of dysplasia tissue relative to normal has been reported [44]. Various excitation light wavelengths (i.e., 330~340, 350~380 and 400~450 nm) were optimized for fluorescence spectroscopic diagnosis of cervical dysplasia [29]. Benavides *et al*. developed multispectral digital colposcopy (MDC) to measure multispectral autofluorescence and reflectance images of the cervix by using an inexpensive color Charged Couple Devise (CCD) camera for *in vivo* detection of cervical dysplasia [45]. Another pilot study demonstrated that Multispectral Digital Colposcopy has the potential for *in vivo* detection of cervical intraepithelial neoplasia [46]. Apart from the above mentioned reports, there is other work reported on fluorescence spectroscopic diagnosis of cervical precancer and cancer [47-50].

1.2.2 Diffuse reflectance spectroscopy

Reflectance spectroscopy is sensitive to the absorption and scattering properties of tissues [51]. Studies by Mirabal *et al*., Skala *et al*. and Marín *et al*. on *in vivo* reflectance spectroscopy to detect cervical precancer have been reported [52-54]. Reflectance spectroscopy demonstrated an inferior diagnostic performance with respect to fluorescence spectroscopy [55]. However, attempts to improve the cervical diagnosis combinations of reflectance spectroscopy with other spectroscopic techniques have been carried out. Nordstrom *et al*., in 2001

characterized cervical precancer and benign changes using fluorescence and diffuse reflectance spectroscopy [56]. Chang *et al*. corroborated the superiority of fluorescence spectroscopy over reflectance and demonstrated the improved diagnosis by a combination of reflectance and fluorescence [55]. Weber *et al*. also used combined reflectance and fluorescence model to fit the *in vivo* measured spectra of cervix [57]. Georgakoudi *et al*. studied the potential of 3 spectroscopic techniques (intrinsic fluorescence, diffuse reflectance, and light scattering) individually and in combination (trimodal spectroscopy) for the detection of cervical dysplasia [58]. The authors concluded that trimodal spectroscopy combining intrinsic fluorescence, diffuse reflectance, and light scattering detects intraepithelial lesions more effectively than any of these techniques alone.

1.2.3 Infrared absorption spectroscopy

Pioneering work was carried out by Wong *et al*. on the utility of Infrared (IR) absorption spectroscopy for detection of dysplasia in cervical cells and tissues [59, 60]. The molecular, structural and biochemical changes associated with cervical dysplasia, such as reduction in glycogen as well as hydrogen bonding of C-OH groups of carbohydrates and proteins, increased degree of disorder of methylene chains of membrane lipids, extensive hydrogen bonding of the phosphodiester groups of nucleic acids, decreased ratio of methyl-to-methylene, red-shift of IR band at 1082 cm^{-1}, and hydrogen-bond strength amide groups decreased in α-helical segments but increased in β-sheet segments, have been observed. Principal Component Analysis (PCA) to discriminate the IR spectra of 272 exfoliated cervical cell specimens was utilized by Wood *et al*. and they

observed that 86% IR-predicted normal cells showed normal pap smear whereas 71% of the IR-predicted malignancies were confirmed by biopsies [61]. Cohenford *et al*. carried out a similar kind of study on a larger cohort of 436 patients with 79% sensitivity and 77% specificity in the cervical cancer diagnosis [62].

Improved sensitivity of 98.6% and 98.8% specificity for cervical dysplasia detection was reported by Fung *et al*. [63]. Subsequent studies by Chiriboga *et al*. on cervical tissues and exfoliated cells, led to a comprehensive understanding of the IR spectral profile [64-66] . This group also correlated spectra of exfoliated cells to the spectra acquired from various layers of cervix tissue sections for better understanding. Based on IR band intensity, Wood and co workers attempted to reveal the potential confounding variables, which could hinder or weaken the discrimination of neoplastic cells [67]. The findings of their work revealed that variables like saline, leukocytes, C. albicans infection, fibroblast, endocervical mucins, sperm contamination and thrombocytes could influence IR spectra. Cohenford and co-workers found that cytologically normal cells from cancer patients are IR spectroscopically different from the normal cells of healthy cases [68]. The study by Romeo *et al*. suggested that endocervical cells as well as benign changes also influence the cervical cancer diagnosis and hence must be removed prior to PCA-based analysis [69]. The same group also established that hormonal status leads to variation in cells spectral features but it does not influence the PCA results [70]. This group also explored the removal of blood components like red and white cells from the cervical smear and their affect on

discriminations [71]. They observed that classification of dysplasia from normal samples became poorer after leukocyte removal.

Apart from the spectral analysis of cell pellets, imaging and mapping of cervix tissues were also carried out. Chang *et al*. recorded the IR images of cervix tissue sections with the ratio intensity of 1130~1180 cm^{-1} and 1180~1260 cm^{-1} IR bands leading to discrimination between normal and dysplastic cervix [72]. Subsequently, to differentiate between tissue pathologies and diverse cell types using IR spectra features, scientists have used cluster analysis to build IR images. Bambery and colleagues attempted to map tissue cryo-sections by both IR band intensity and cluster analysis using unsupervised Hierarchical Cluster Analysis (HCA) [73]. They observed that the similar layers were grouped into the equivalent clusters. Another study by Wood *et al*. successfully identified images of normal, Low grade squamous intraepithelial lesions (LSILs) and High grade squamous intraepithelial lesions-(HSILs) tissues with a similar method and discriminated dysplasia from normal tissues [74]. IR spectroscopy, coupled with fuzzy C-means clustering for data reduction and HCA for classification have the ability to distinguish different tissue pathology types [75]. A neural network was introduced for developing cervical diagnostic algorithms to optimize the performance of IR spectra for screening cervical dysplasia. This aids in grading cervical intraepithelial neoplasia (i.e., CIN 1, 2 and 3) with the use of probabilistic neural networks (PNN) with an accuracy rate of ~85% [76]. A subsequent study by the same group, differentiated cancer from normal tissues with an improved accuracy of >95 [77].

1.3 Raman spectroscopy in cervical cancer diagnosis

1.3.1 Raman effect

Sir C. V. Raman first observed Raman scattering in 1928. This phenomenon is based on inelastic scattering of light [78]. He used a very simple system to observe the Raman scattering, which included filtered sunlight as an excitation source, a liquid as sample, a colored filter as a monochromator; a human eye and subsequently the photographic plate as detector. The fundamental concept of a Raman spectroscope remained similar until technological advancements have made more sophisticated individual components. It was the development of the laser in 1960s and CCD in 1980s that opened up a completely novel area of optical and spectroscopic research [79].

When a photon interacts with a molecule, a majority of scattered photons exhibit no frequency changes relative to the incident photons a phenomenon called an elastically scattering. This phenomenon is also called Rayleigh scattering. This is a classical theory of light scattering proposed by Lord Rayleigh in 1871. A small fraction of photons (approximately 1 in 10^8) undergo an energy exchange with the molecule with a resultant shift in frequency, as compared to incident photons. The process is called an inelastic scattering or Raman scattering. Raman scattering is a rare event [80]. According to the light scattering theory, the interaction of light with a molecule leads to a polarization of the molecule and then the polarized molecule shows an induced dipole moment caused by the external field [81]. The induced dipole moment, P, is directly proportional to the

electric field E and to a property of the molecule called the polarizability α, as shown in the following equation.

$P = \alpha E;$ $E = E_0 \cos 2\pi \nu_0 t;$ $P = \alpha E_0 \cos 2\pi \nu_0 t \ldots\ldots$ (Equation 1.1)

where E_0 is electric field amplitude and ν_0 is the frequency of the incident light. The polarizability α is dependent upon the instantaneous positioning of molecule nuclei. For a molecule having N atoms, it has 3N degrees of freedom. Of these, 3N-6 (3N-5 for a linear molecule) results in vibrations of the molecule. The induced dipole moment for a diatomic molecule with a single normal coordinate is as follows [82].

$$P = \alpha E_0 \cos 2\pi \nu 0 t + \frac{1}{2} \mathrm{E_0 Q_1^0} \left(\frac{\partial \alpha}{\partial \mathrm{Q_1}}\right)_0 A = \pi r^2 \times [\cos 2\pi \mathrm{t}\,(\nu 0 + \nu 1) + \cos 2\pi \mathrm{t}\,(\nu 0 - \nu 1)]$$

…….…….. (Equation 1.2)

where α_0 is the intrinsic polarizability of the molecule, Q^o_1 is the vibrational amplitude and ν_1 frequency of the vibration. The first term of equation 1.2 represents Rayleigh scattering and if is nonzero then Raman scattering occurs. The second and third terms represent anti-Stokes and Stokes Raman scattering, respectively. When the scattered photon has a lower energy than the exciting photon then it is called Stokes Raman scattering, whereas, if the scattered photon is shifted to shorter wavelength compared to the exciting photon, is called an anti-Stokes Raman scattered photon [82].

As per the Quantum theory of Raman effect, radiation is a stream of particles called photons with a frequency 'ν' having energy '$h\nu$' where 'h' is a Planck's constant [83]. One of the several phenomena that can occur upon

interaction between quantum of radiation $h\nu_{ex}$ and molecule is scattering. Thus the incident quantum of photons may cause an elastic scattering with unchanged ν_{ex}, the Rayleigh scattering (Figure 1.3). The total energy of the molecule and photon is conserved.

This phenomenon can be explained as

$h\nu_{ex} + W_1 = h\nu_m + W_2$……. (Equation 1.3)

where,

$h\nu_m$ is the energy difference between two vibrational molecule's energy states

W_1 is the energy of the molecule before the interaction

W_2 is the energy of the molecule after the interaction

If $W_1 > W_2$, then $\Delta\nu$ is positive; and an anti-strokes line results. If $W_1 < W_2$ then $\Delta\nu$ is negative; and strokes line results.

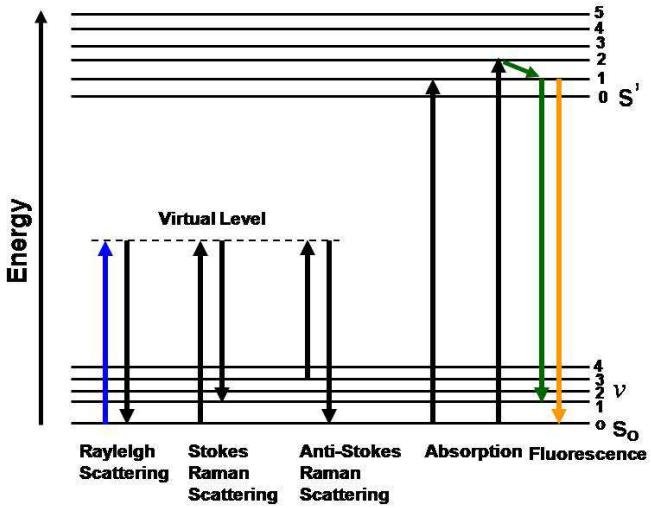


Figure 1.3: Energy transition diagram of vibrational spectroscopy. *v* is the vibrational quantum number.

(http://www.gamry.com/application-notes/EIS/raman-spectroelectrochemistry/)

It is obvious that most often the anti-strokes line originate due to interaction between photons and molecules in excited quantum states. Assuming that the

scattering medium is in temperature equilibrium at temperature T, the distribution of the molecules over the energy states will be Bolzmann and the ratio of the number of molecules N_1 in a state of energy W_1 to the number of molecules N_2 in a state of energy W_2 is given by

$$\frac{N_1}{N_2} = e^{-h\Delta v/kT} \quad \text{(Equation 1.4)}$$

If one includes the fact that the scattered intensity is proportional to the fourth power of the frequency then the relative intensities of stokes to anti-Stokes and their temperature dependence should be consistent with the ratio [83].

$$\frac{I_{anti-Stokes}}{I_{Stokes}} = \left(\frac{v+\Delta v}{v-\Delta v}\right)^4 e^{-h\Delta v/kT} \quad \text{(Equation 1.5)}$$

Hence, as described by above Bolzmann's equation, the exponential term is dominant; this explains the weak nature of anti-Stokes as compare to Stokes line.

Raman shift

The arithmetical energy difference between the initial and final vibrational levels (v) or Raman shift in wavenumber (cm^{-1}) can be calculated by the following equation.

$$v = \frac{1}{\lambda_{incident}} - \frac{1}{\lambda_{scattered}} \quad \text{(Equation 1.6)}$$

where 1/λ incident and l/λ scattered are the frequencies in cm^{-1} for the incident and Raman scattered photons, respectively. The differences in energy between the incident photon and the Raman scattered photon is equivalent to the vibrational energy of the scattering molecule. The intensity plot of scattered light

against energy difference between incident and scattered photons i.e. Raman shift is called a Raman spectrum. As the intensity of Raman scattering is low, the heat dissipation does not lead to a computable temperature increase in the system [84]. The specific group of vibrational bonds in the molecules of the specimen gives distinct bands, which are characteristic of the Raman spectrum. Raman spectroscopy has been known to have potential for various cancer and precancer diagnosis, based on Raman spectral variations representing the bio-molecular, structural and conformational changes associated with tumorogenesis.

In addition to Raman scattering, there exists another spectroscopic phenomenon called absorption (Figure 1.3). The amount of energy exchange during Raman scattering is equal to the energy absorbed in IR absorption. The frequency shift for a specific vibration band of the same molecule remains the same for Raman scattering and IR absorption. However, the selection rules differ for IR and Raman scattering. When a dipole moment changes during the molecular vibration then a molecule can absorb IR light. The Raman effect occurs by an oscillation-induced dipole moment. This means that the molecular interaction with the photon is through the polarizability of the molecule. Thus, all the molecules are not Raman-active and IR-active, which makes Raman spectroscopy and IR spectroscopy complementary to each other.

In recent years, different modalities to Raman spectroscopy such as Spatially Offset Raman Spectroscopy (SORS), coherent anti-Stokes Raman Spectroscopy (CARS), Resonance Raman Spectroscopy (RRS), Surface Enhanced Raman Spectroscopy (SERS), etc. are employed in different areas of

science and technology [85, 86]. Spatially offset Raman spectroscopy (SORS) detects the signal offset from the point of laser excitation to allocate below surface measurements, for example, subcutaneously *in vivo* [87]. Coherent anti-Stokes Raman Spectroscopy (CARS) allows vibrational imaging with high sensitivity, elevated spectral resolution and three dimensional sectioning potential. It relies on inducing signal in the target molecule using two lasers, probed by a third laser which creating a coherent signal in the phase-matching direction at a blue-shifted frequency [88]. In Resonance Raman spectroscopy (RRS), Resonance Raman spectra are obtained when the energy of photon of an exciting laser beam matches or is close to the energy require for electronic transition [89]. Surface-enhanced Raman scattering (SERS) takes advantage thousand to 10 million fold effective increase in signal for molecules attached to or near nanometer-sized metallic structures (gold/silver colloids), enabling single molecule level detection [88].

1.3.2 Raman instrumentation

A conventional Raman spectrometer consists mainly of four components, including an excitation light source, filters, spectrograph and detector. In principle, a beam of photons hits a specimen and scattered photons are collected, filtered and dispersed and detected to generate a Raman spectrum. A graphic illustrating a typical Raman instrument is shown in Figure. 1.4. Concise descriptions of the individual components are presented in the following sections.

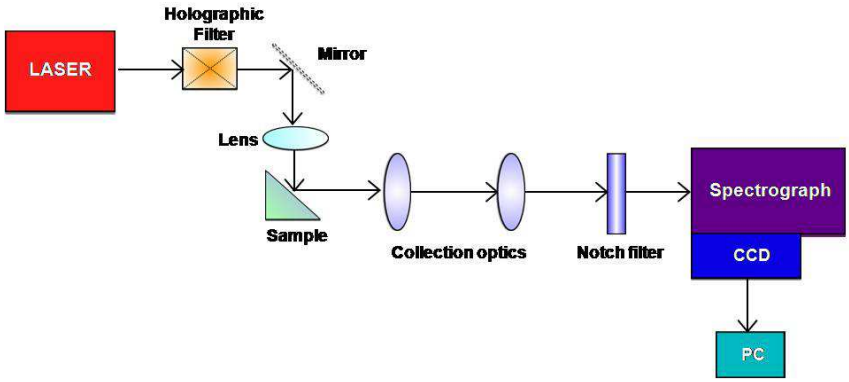


Figure 1.4: Graphic representation of a typical Raman spectroscope.

1.3.2.1 Excitation Source (LASER)

The advent of laser as a powerful and monochromatic source of excitation light improved the Raman signals considerably [90]. Usually, lasers are the light source in Raman spectroscopy because of their higher power output and narrow bandwidth. The choice of wavelength for Raman measurements depends on the specific applications and the spectroscopic properties of the specimen. For biomedical applications, a Near Infrared (NIR) laser is commonly used due to its penetration depth and low level of tissue autofluorescence [91]. Generally, a laser system consists of a lasing medium (atom, molecule or ion), a resonant cavity and an excitation source (electrical, radiation or others). The excitation source excites the laser medium's atoms/ions to a higher energy level. The transition from excited state to a lower state produces the laser radiation, and this is amplified by stimulated emission because of single/multiple passes through the resonant cavity. Lasers can be categorized into types like solid, liquid, gas and diode lasers. Solid-state and external-cavity-stabilized diode lasers are popular choices in biomedicine due to their portability. The current real-time Raman system utilized

in the study is equipped with a wavelength-stabilized diode laser (785nm). Additional details about this laser have been provided in chapter 2, under the methodology section.

1.3.2.2 Filters

Filters are utilized to filter out the intense Rayleigh scatter or unwanted background signals. Broadly, filters are classified into two types named band-pass and edge filters. Edge filters are further categorized into Long-pass and Short-pass filter. An optical interference Long-pass (LP) filter attenuates shorter wavelengths and transmits longer wavelengths over the active range of the target spectrum (ultraviolet, visible, or infrared). Long-pass filters can have a very sharp slope and are named according to the cut-off wavelength at 50 percent of peak transmission. A short-pass (SP) filter is an optical interference or else coloured glass filter that attenuates longer wavelengths and allows shorter wavelengths above the active range of the target spectrum (usually the ultraviolet and visible region). Band-pass filters transmit a specific range of wavelength and obstruct the other wavelengths. The spectral width of such a filter is expressed by the wavelength range which it lets through and can be anything from less than Angstroms to a few hundred nanometers. Band-pass filters can be prepared by uniting a LP and SP filter.

1.3.2.3 Spectrograph

The key function of the spectrograph is to separate the light from an object into its constituent wavelengths. A spectrograph consists of four essential components: an entrance slit, a collimating element this can be a lens/mirror to obtain parallel rays when pass through the entrance slit, a dispersing element-the key component.

Generally, a grating utilized to spread the light in space as a function of wavelength and a focusing element to form an image of the entrance slit at the detector focal plane. If monochromatic light is incident on a grating surface, it is diffracted into distinct direction. The light diffracted by each groove combines to form set of diffracted wave fronts.

1.3.2.4 Detector

The detector acquires the intensity of Raman signal at each wavelength. A charged-coupled detector (CCD) can be pictured as an array of photosensitive facets or pixels. In Raman spectroscopic applications, the wavelength/Raman shift corresponds to the horizontal rows. Moreover, the column pixels are typically binned vertically, providing the intensity at each wavelength. The CCD is fabricated on a silicon chip, typically of 1024 × 256 pixels. Each of these is 25 μm on a side, and the array covers an area of about 25 mm × 6 mm. Due to advancements in the CCD technologies, quantum efficiencies of 90% can be achieved.

1.4 Data analysis

As mentioned previously, histopathology is the gold standard for cervical cancer diagnosis. However, it suffers from subjectivity, as it involves careful visual inspection of the suspected section of the tissue under the microscope by an experienced pathologist. Fatigue factors due to examination of large number of samples and inexperience have been reported to worsen the error rate in the conventional approach of cancer diagnosis. These problems can be overcome in spectroscopic diagnosis. An important aspect of optical spectroscopy is

objectivity. As spectral data are amenable towards statistical tools, the objectivity can be achieved. The applications of these tools facilitate the computation of mathematical parameters derived from spectral data for classification. The distinct feature of this approach is that it is devoid of visual decision making and the system (computer) is completely blind to the sample that is being analyzed. Analysis of the data generated from a spectroscopy experiment can be performed in two different ways: univariate and multivariate. Univariate analysis using optical density values is generally performed in colorimetric estimations of different biomolecules [92]. In this case, knowing either of two dependent and independent variables, a solution for the second variable can be calculated. In contrast to the univariate approach, multivariate analysis involves observation and analysis of more than one statistical variable at a time [92]. Data generated from infrared or Raman experiments consist of results of observations of multiple variables (wavenumbers) for a number of individuals (diseased or healthy). Each variable may be regarded as constituting a different dimension, such that if there are ‘n’ variables (IR or Raman bands). Each object may be said to reside at a unique position in an abstract entity referred to as n-dimensional hyperspace. This hyperspace is necessarily difficult to visualize. The underlying theme of multivariate analysis (MVA) is simplification or dimensionality reduction. This can occur in one of two ways; either using an unsupervised or a supervised learning algorithm. In general, unsupervised methods such as principal component analysis (PCA) and hierarchical cluster analysis (HCA) are used to assess the ‘natural’ differences and similarities between spectra. These methods

are employed to discover structure in the data and can be used to 'cluster' samples into groups by producing scatter plots (PCA) and dendrograms (tree-like figures; HCA). By contrast, supervised methods like linear discriminant analysis (LDA) and artificial neural networks (ANNs) are 'calibrated' with some known existing parameters about the sample. A prior knowledge is used in the construction of the LDA or ANN model followed by validation of model with test data or cross validation [93-96]. In the following sections, a brief description of multivariate data analysis methods employed in the thesis is presented.

Principal Component Analysis (PCA)

Principal component analysis (PCA) is an unsupervised multivariate method for data compression and over-viewing. Arithmetically, it is an eigenvector decomposition of the variable correlation matrix. It recognizes trends, pattern and outliers in the data set [96, 97]. It decomposes the data into their most common variation (factors) and produces small set of well defined numbers (scores) for each sample that represent the amount of variation. Schematic illustration of working principle of PCA is shown in Figure 1.5.

The representative matrix model for PCA can be expressed by the following equation

$$A = S\,F + EA \qquad \text{(Equation 1.7)}$$

where

A is an n by P matrix of spectral absorbance,

S is an n by f matrix of score values for all of the spectra,

F is an *f* by *P* matrix of eigenvectors, and

EA matrix is the residual spectra matrix

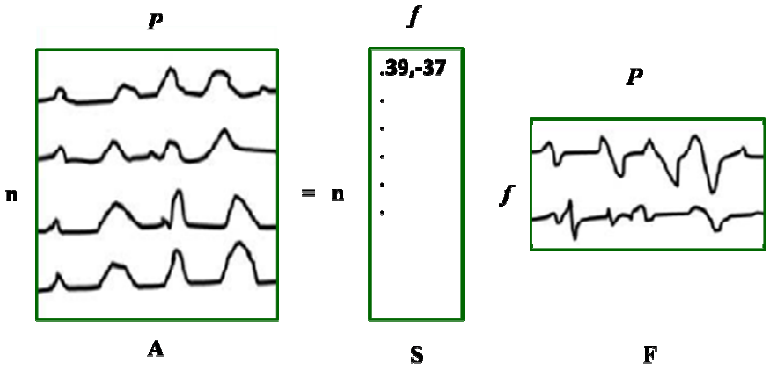


Figure 1.5: Schematic illustration of working principle of PCA, A= original spectral data; n = number of spectra; S = PCA scores, P = number of data points; F = PCA factors (Eigenvectors, Loadings); f = number of principal components.

In Figure 1.5, *n* is the number of samples (spectra), *P* is the number of data points (wavelengths) used for calibration, and *f* is the number PCA eigenvectors (PCs). Usually PCs are calculated by two techniques; the Non-linear Iterative Partial Least Squares (NIPALS) algorithm or Decomposition of covariance [98]. The PCs (eigenvectors) are orthogonal, maximize the data variance and they corresponds to their respective eigenvalues [98]. The first PC describes the greatest variance from the mean while the last PC with lest variance and the smallest eigenvalue. First few PCs will show significant variations and are included in analysis. This makes the PCA well suited for multivariate data visualization and interpretation. PCA aims to summarize the overall variability, which includes both the inter-group divergence, and the intra-group variation. However, to assess the relationship between the different clusters, an adequate

method should focus on inter-group variability, while neglecting intra-group variation.

Principal Component-Linear Discriminant Analysis (PC-LDA)

In PC-LDA, PCA is first carried out on the entire data set to reduce the dimensionality of the data while preserving the diagnostically significant information for classification. As mentioned earlier, PCA describes data variance by identifying a new set of orthogonal features, called as principal components (PCs) or eigenvectors. Due to their orthogonal characteristics, the first few PCs are sufficient to represent maximum data variance. Every eigenvector is associated with the original spectrum by a variable named the PC score, which characterizes the rank of that particular component against the source spectrum. Differences between different classes are reflected by PC scores. The unpaired Student's t-test is used to identify diagnostically significant PCs ($p<0.05$) [99, 100]. These PC scores are then used as input data for LDA based classifications. Although PCA aims to identify features that represent variance in the data; LDA provides data classification is based on an optimized criterion which is objective for more class separability. LDA transformation matrices are generated and it further identifies eigenvectors or LDA components of this classification criterion [100]. The scree plots depict the variance (% correct classifications) accounted for by the total number of LDA components selected for analysis. The outcomes of PC-LDA are generally represented in the form of a confusion matrix, where diagonal elements are true positive and non-diagonal elements are false positive predictions. The confusion matrix aids to understand the separation within the

groups which is acquired by accounting for the contribution of all selected factors used for analysis. PC-LDA results are also represented in the form of scatter plots, generated by plotting various combinations of scores of factors. The best method for developing and validating the efficacy of any diagnostic model is to validate results with an independent test data. An algorithm is fit to the data in the training set using the empirical or statistical method of choice, and the criteria for classification into specific categories is determined. Classification of the spectra in the test set determines the impartial accuracy of the algorithm. However, in cases of small data sets (as is often the case in pilot studies), dissection of the data into training and test sets is not possible. The leave-one-out cross-validation method is a popular alternative to independent test sets. In the leave-one-out method, one spectrum is removed from the data set and the algorithm is driven using the remaining data. The algorithm is then tested using the removed spectrum. This process is repeated for every spectrum in the data set, such that an estimate of the potential accuracy of future algorithms developed using the method in question can be calculated [101]. In the present study, test prediction was also used to validate the models. Algorithms for these analyses were implemented in MATLAB (Mathworks Inc.) based in-house software [101].

1.5 Applications of Raman spectroscopy in cervical cancers

Raman spectroscopy (RS) represents a technique capable of label-free and nondestructively probing endogenous biomolecules (e.g., proteins, lipids, carbohydrates and nucleic acids) to determine highly specific diagnostic information [102]. Each molecule has a unique Raman spectrum at well-defined

frequencies. As normal and abnormal specimens are expected to have differences in biomolecular composition, thus can probe biomolecular changes. Biomedical applications of Raman spectroscopy in numerous organs like cervix, skin, urinary bladder, oral, lung, colon, brain and gastrointestinal cancer have been studied [3, 103, 104]. The pioneering *in vitro* and *in vivo* Raman spectroscopic applications of cervical cancers were reported by Mahadevan-Jansen *et al*. in 1998 [105, 106]. Subsequent *in vivo* studies have demonstrated the feasibility of fast and non-invasive detection of cervical dysplasia [107-116]. Detection of effects due to HPV in cervical cancer cell lines and clinical specimens has also been reported [117-119]. Since 1998, Raman spectroscopy has continued to advance screening/ diagnosis/ treatment prediction response/ applications in cervical cancers. The following section provides an overview of biomedical application of RS in cervical cancer detection.

1.5.1 Cervical cell lines studies

High risk Human papillomavirus (HR-HPV) is well known etiological factor of cervical cancers. Hence, identification of HR-HPV presence has clinical significance. In 2007, Raman spectroscopic identification of the HR-HPV 16 virus related cellular effect was reported by Jess *et al*. They acquired Raman spectra from primary human keratinocytes (PHK), E7 gene expressing PHK (PHK E7) and CaSki cells (HR-HPV 16 containing cervical cancer cell line). Their study showed variations, mostly in peaks originating from DNA and proteins, consistent with HPV gene expression and cellular changes associated with neoplasia [117]. They observed that RS can discriminate between normal

keratinocytes and keratinocytes expressing HR-HPV 16 E7, which extends keratinocyte life span and is sufficient to immortalize these cells, with 93% sensitivity and 93% specificity. They concluded that Raman microspectroscopy can identify cells expressing the HR-HPV 16 E7 gene accurately and objectively, suggesting that this approach may be of value for the identification and discrimination of the different stages of HR-HPV associated neoplasia. In 2010, a subsequent study by Ostrowska *et al*. aimed to investigate biochemical changes in cells caused by high-risk HR-HPV strands (HPV-16 and HPV-18). They also investigated differences between the cells with high, medium and low HR-HPV copy number, using vibrational spectroscopic techniques [118]. In this study, Raman and Fourier transform infra-red absorption (FTIR) spectra were acquired and investigation of four cervical cancer cell lines: HPV negative C33A, HPV-18 positive HeLa with 20–50 integrated HPV copies per cell, HPV-16 positive SiHa with 1–2 integrated HPV strands per cell and HPV-16 positive CaSki containing 60–600 integrated HPV copies per cell. They observed that vibrational spectroscopic techniques can discriminate between the cell lines and elucidate cellular differences originating from proteins, nucleic acids and lipids. The study by Vargis *et al*. evaluated the ability of RS to detect the presence of HPV and differences between specific HPV strains [119]. In this study, two sets of experiments were conducted to determine the sensitivity of RS in detection of HPV infection. First, Raman spectra were acquired using a Raman confocal microscope from four different cell lines: HPV-16–positive SiHa cells, HPV- 18–positive HeLa cells, HPV-negative but malignant C33A cells, and benign NHEK

cells. Next, Raman spectra were obtained from HPV-positive and HPV-negative patient samples. Their study showed that spectra from the cell culture lines and the patient samples contained many statistically significant differences. As cell lines are maintained in different environmental conditions, cell lines are expected to have differences in their biochemical constituents from those of patient specimens. They concluded that Raman microspectroscopy can be used to detect HPV and differentiate specific HPV strains. The studies showed the promise and many have similarities in terms of high risk HPV detection. However, the studies in this area need to focus on probing HPV induced neoplastic changes in the cells. It is also important to understand that studies on single cells are important to get the correct signature of HPV infected and neoplastic transformation of cell, but Raman studies on cell pellets are equally important with a view forwards rapid diagnosis.

1.5.2 *Ex vivo* studies

Raman spectroscopic studies on *ex vivo* tissue can be carried out either on tissue biopsies (conventional spectroscopy) or on the tissue sections (imaging). Studies using *ex vivo* tissue have shown encouraging results for the application of Raman spectroscopy for improving the detection of cervical cancer. This section will elaborate the Raman spectroscopic *ex vivo* studies on cervical cancer tissues.

Ex vivo Raman spectra of fresh, frozen and preserved cervix tissue biopsies have been reported in the literature [106, 120]. In 1998, primary *ex vivo* study on cervix tissues indicated the potential advantage of Raman spectroscopy for diagnosis of cervical precancers [106]. Raman spectra of 36 cervix tissues from

18 patients were measured. Out of 36 tissues, 19 were normal, 2 were metaplasia, 4 were inflammation, two were HPV positive and nine were precancers. In this study two different algorithms were employed for tissue differentiation. The first method used empirically selected peak intensities and ratios of peak intensities to differentiate precancers from other tissue categories. In the second method authors of the study employed multivariate statistical method to differentiate precancers from other tissues. They concluded that empirically selected normalized intensities can differentiate precancers from other tissues with sensitivity and specificity of 88% and 92%, respectively. However, unbiased multivariate methods gave a sensitivity of 82% and specificity of 92%. These algorithms can potentially separate benign abnormalities such as inflammation and metaplasia from precancers. They also compared tissue spectra to earlier reported spectra and empirically measured chromophore spectra; they reported that collagen, nucleic acids, phospholipids and glucose l-phosphate to be most likely contributors to the Raman spectra. In 2006, the another *ex vivo* study was carried out on tumor and normal tissue [120]. They reported that Raman spectra of normal cervix tissues were characterized by strong broad amide I, broader amide III and strong peaks at 853 and 938 cm^{-1}, which were attributed to structural proteins such as collagen. Prominent features of malignant tissue spectra, with respect to normal tissue were relatively weaker and sharper amide I, minor red shift in δCH_2 and sharper amide III which indicated the presence of deoxyribonucleic acid (DNA), lipids and non-collagenous proteins. PCA

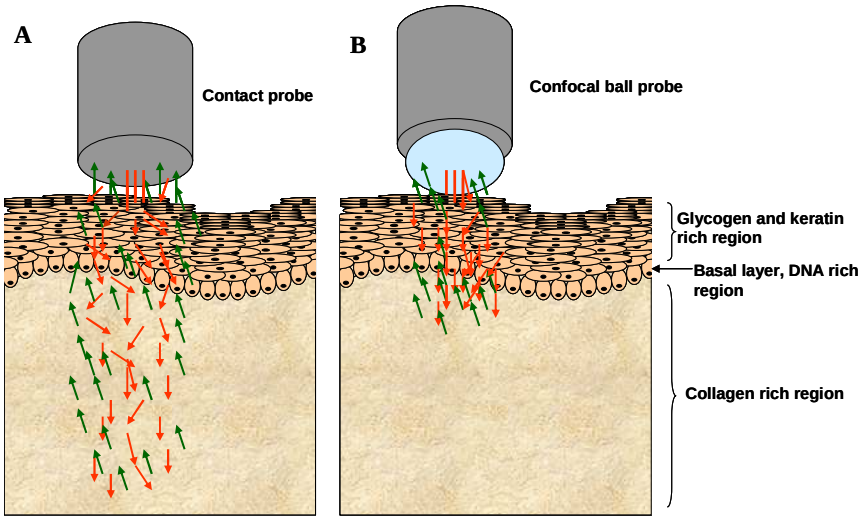
combined with a multiparametric limit test was used for discriminating normal and cervical cancers with 99.5% sensitivity and specificity.

Raman imaging on tissue sections is far less commonly reported, this could be due to the extremely long acquisition time for pixel-by-pixel acquisition. However, recent technological advancements in Raman imaging have reduced the acquisition time greatly [121]. Previous studies on tissue sections included spectral acquisition from different layers of tissue. A study by Faoláin *et al*. in 2005 directly compared Raman spectroscopy and synchrotron infrared (SR-IR) spectroscopy on parallel cervical cancer samples [122]. In their study, they used frozen and dewaxed formalin paraffin preserved tissue and could discriminate between different cell types in normal cervical tissue. The spectra of invasive carcinoma showed marked differences from normal cervical epithelial cells. They observed that spectral differences associated with the onset of carcinogenesis include increased nucleic acid contributions and decreased glycogen levels. A subsequent study in 2007 by Lyng *et al*. investigated the potential of Raman spectroscopy as a diagnostic tool to detect biochemical changes associated to cervical cancer progression [123]. In their study, Raman spectra were acquired from different point of de-waxed 10 μm sections, which were obtained from formalin-fixed paraffin preserved (FFPP) tissue blocks of 20 normal and 20 invasive carcinomas subjects. They also acquired the Raman spectra from pure compounds of proteins, nucleic acids, lipids and carbohydrates in order to gain an insight into the biochemical composition of cells and tissues. In the study, Raman spectra from basal cell, epithelial cell and connective tissue were acquired from

normal cervix tissue. They observed that spectra of basal cells show strong bands at 724, 779 and 1578 cm^{-1} which are characteristic of nucleic acids. Spectra of epithelial cells showed characteristic glycogen bands at 482, 849, 938, 1082 and 1336 cm^{-1}, whereas spectra of connective tissue showed characteristic collagen bands at 850, 940 and 1245 cm^{-1}. They reported that the absence of glycogen bands, the presence of characteristic nucleic acid band and an increase in the intensity of the amide I band was observed in the spectra from invasive carcinoma. Spectral features observed in invasive carcinoma specimens were also observed in the premalignant specimens such as the nucleic acid bands at 724, 779, 852, 1366 and 1578 cm^{-1}. However, these studies did not look into the spectral differences in basal cells to that of tumor. It is vital to note that abnormality in the cells of the basal layer develop into the neoplastic tumor cells. Hence, it is important to probe the differences between basal cells and tumor cells, as both the cells are in the dividing phase but the proliferative index of tumor cells is high as compared to the basal cells. In 2008, Martinho *et al*. studied the Raman-based optical diagnosis of normal cervix, inflammatory cervix (cervicitis), and cervical intra epithelial neoplasia (CIN) with 63 specimens [124]. They found the main alterations in the 857 cm^{-1} (CCH deformation aromatic); 925 cm^{-1} (C-C stretching); 1247 cm^{-1} (CN stretch, NH bending of Amide III); 1370 cm^{-1} (CH2 bending); and 1525 cm^{-1} (C=C=C=N stretching) vibrational bands. In 2010, Kamemoto *et al*. reported *ex vivo* microRaman spectroscopy study on normal and cancerous cervical human tissue section from 7 patients [125]. They observed the spectral features associated with collagen (775 to 975 cm^{-1}) in

normal squamous cells which were below the detection limit in cancer. In their study, Raman chemical maps of regions of cancer and normal cells in the cervical epithelium made from the spectral features in the 775 to 975 cm^{-1} and 2800 to 3100 cm^{-1} regions were generated. The authors' interpretation of presences of collagen in normal squamous cells did not match 'earlier reports. However, earlier reports suggested the presence of collagenous protein features in connective tissue and glycogen in epithelium i.e. normal squamous cells [123]. Studies of *ex vivo* cervix tissue biopsies showed similarity in spectral assignments, whereas studies on cervix tissue sections exhibited discrepancy in the spectral assignments.

It is very important to understand this discrepancy. Differences between these reports may be due to the comparison between different layers of normal tissue to that of tumor. The architecture of tissue consist of 8-10 layers of cells called the epithelium, resting on connective tissue which is rich in collagen, as shown in Figure 1.5., while the epithelium is very thin, consisting of the basal layer, intermediate layer and superficial layer.



1.6: Schematic representation of cervix tissue model for different probing area (A) Contact probe (B) Confocal probe.

Raman signals depend on the probing volume, the kinds of cells, and the part of tissue being probed. For example, in Figure 1.5 A, assume a contact probe acquiring the spectra from one millimeter depth of penetrance, in this case Raman photons from the stomal region will dominate and hence spectra will exhibit collagenous rich signals in the Raman spectra. Alternatively in Figure 1.5 B, imagine the confocal probe measuring Raman signals from the few micrometers of epithelial region and minimal signals from the stromal region. In this case, spectra will be rich in glycogen signals.

1.5.3 *In vivo* studies

A number of *in vivo* Raman studies has been performed, in which a portable Raman spectroscopic fiber-probe system was utilized for diagnosis of different cancers including cervical cancers [79, 106-108, 110-116, 126]. In 1998, Mahavedan *et al*. developed a fiber optic probe to measure *in vivo* Raman spectra

of cervix to detect precancer lesions from the normal cervical tissue [106]. In their study, they could acquire Raman spectra in 90 seconds from normal and precancer cervix tissues *in vivo*. Their study also concluded that increasing the power of the excitation source could reduce integration time to below 20s, enabling the measurement of Raman spectra in cervical tissues. Moreover, another pilot clinical trial on 13 subjects by Utzinger *et al*. was carried out by measuring Raman spectra of precancer lesions and corresponding normal cervical tissues, suggesting that *in vivo* Raman spectra resemble those of in vitro cervix tissues [107]. Their studies also concluded that cervical epithelial cells may contribute to tissue spectra at 1330 cm^{-1}, a region associated with DNA and epithelial cells probably do not contribute to tissue spectra at 1454 cm^{-1}, a region associated with collagen and phospholipids. The study conducted by Robichaux-Viehoever *et al*., in the year 2007, on 79 subjects using a clinical feasible time (5 seconds), indicated that Raman spectroscopy can distinguish between high-grade dysplasia and benign tissue with sensitivity and specificity of 89% and 81%, respectively [108]. The *in vivo* Raman study by Kanter *et al*., on 90 subjects to differentiate between normal ectocervix, normal endocervix, low grade dysplasia and high-grade dysplasia suggested that Raman spectroscopy in conjunction with the diagnostic algorithm can distinguish dysplasia from normal ectocervix with classification efficiency of 95% [111]. Moreover, their study showed that Raman spectroscopy can differentiate between different precancers with improved sensitivity of 98% and specificity of 96% by using multiclass discrimination algorithms like maximum representation and discrimination feature (MRDF) and

sparse multinomial logistic regression (SMLR). In the year 2009, the same group come up with an interesting finding, on 120 subjects, by acquiring *in vivo* Raman spectra from normal, low-grade metaplasia, high-grade metaplasia and metaplasia cervical tissues [111]. They incorporated a hormonal status parameter, specifically the point in the menstrual cycle and menopausal status, and the classification accuracy of their algorithm improved from 88% to 94%. Moreover, in the same year, they also reported data from 113 subjects stratified by menopausal state which resulted in an improvement of the accuracy of classification of low grade squamous intraepithelial lesion (LSIL) to 97% from 74% [110]. This concludes that RS is almost one step closer to clinical use by simply improving sensitivity to differentiate LGSIL from normal. Another study in the same year by Mo *et al*. reported that the high wavenumber region of Raman spectra can be used for diagnosis and detection of precancer cervix [112]. Their study showed that the diagnostic algorithms based on principal components analysis and linear discriminant analysis together with the leave-one patient-out cross-validation method yielded a diagnostic sensitivity of 93.5% and specificity of 97.8% for dysplasia tissue identification. The results of the study by Duraipandian *et al*. on 29 subjects, suggested RS in conjunction with genetic algorithm-partial least squares-discriminant analysis (GA-PLS-DA) with double cross-validation (dCV) methods has the potential to provide clinically significant discrimination between normal and precancer cervical tissues at the molecular level [113]. Vargis *et al*., in 2011, reported sensitivity of Raman spectroscopy to normal patient variability such as race/ethnicity, body mass index (BMI), parity,

and socioeconomic status [114]. Their results suggest that BMI and parity have greatest impact whereas race/ethnicity and socioeconomic status have a limited effect. Subsequently, in 2012, another study on 44 subjects, by the same group reported that simultaneous fingerprint/high wavenumber confocal RS has the potential for early diagnosis and detection of cervical precancer *in vivo* [115]. The authors successfully developed an integrated fingerprint (FP) and high-wavenumber (HW) Raman signals of a cervix *in vivo* with 85% sensitivity and 81% specificity. The authors also reported that spectral differences between normal and dysplastic cervical tissue were related to protein, lipids, glycogen, nucleic acids and water content in tissue. The same group in 2013 reported a study on 84 subjects using confocal RS that confocal *in vivo* RS has great potential to improve early diagnosis of cervical precancers [116]. Confocal Raman spectroscopy coupled with PC-LDA modeling yielded a sensitivity and specificity of 81.0% and 87.1% respectively, for *in vivo* discrimination of dysplastic cervix. *In vivo* studies of Raman spectroscopy in cervical cancer diagnosis have been listed in Table1.1.

The laser power used in earlier studies was very low with longer spectral acquisition time but in recent years, increases in laser power and decrease in spectral acquisition time suggests the improvement in technology and it clinical applicability.

Table 1.1: List of *in vivo* Raman studies in cervical cancer detection.

Ref. No.	Acquisition time (s)	Laser power (mW)	No. of cases	Major findings
106	90	15	----	Raman spectra (RS) can be measured *in vivo* from cervix
107	60-180	15	13	*Ex vivo* and *in vivo* RS show similarities
109	5	80	79	Classification of high-grade squamous dysplasia and normal
110	5	80	90	Classification of ectocervix, endocervix, low grade and high grade dysplasia shows multiclass algorithm is better for classifications
111	3	80	120	Incorporation of hormonal status improves classification
114	3	80	133	Incorporation of menopausal status improves classification
112	1	100	46	High-wavenumber can detect cervical dysplasia
114	2-3	80	75	Body mass index and parity have the greatest impact on classification
112	1	100	29	Genetic algorithm-partial least squares-discriminant analysis (GA-PLS-DA) with double cross validation identify cervical dysplasia
115	1	100	44	Simultaneous fingerprint and high wavenumber has potential to detect cervical dysplasia
116	1	100	84	Confocal Raman spectroscopy has potential to improve early diagnosis.

Aims and objectives

In vivo Raman spectroscopy has demonstrated the feasibility to detect cervical dysplasia. Translation of this technology into clinics requires further careful validation on diverse populations and larger cohorts. In addition, no Raman study to classify the cytological certified exfoliated cervical cell specimens has been accounted. The present dissertation aims to evaluate the efficacy of Raman

spectroscopic methods for non-invasive/minimal-invasive and objective screening/diagnosis of cervical cancers under clinical setting.

The thesis focuses on following objectives:

- ✓ To characterize Raman spectral differences between normal and cancerous cervical tissues, in both, *in vivo* and *ex vivo* conditions.
- ✓ To characterize Raman spectral differences between normal, pre-cancerous and cancerous cervical exfoliated cells.
- ✓ To characterize Raman spectral differences in HPV expressing and non expressing cell-lines.

2 *EX VIVO* AND *IN VIVO* RAMAN SPECTROSCOPY ON CERVICAL CANCERS

Introduction

The application of *in vivo* Raman spectroscopy in cervical cancer diagnosis has been studied since 1998; many variables like HPV infections, menopause, hormonal status, race, ethnicity, body mass index, parity, socio economical status, different wavenumber regions, various analytical tools and confocal Raman probes have been explored [27, 108, 110-116]. Despite of large number of studies in this arena, further validations on diverse population and larger cohort are necessary for the translation of this technology into clinics.

The current chapter presents an *in vivo* Raman spectroscopic study for cervical cancer diagnosis in Indian population. This chapter discuss about the standardization of the *in vivo* Raman setup utilizing *ex vivo* cervix tissue specimens and the utility of the vaginal site as an internal control.

2.1 Experimental methods

2.1.1 Raman system utilized in *ex vivo* and *in vivo* studies

Commercial instrument HE-785 (Jobin-Yvon-Horiba, France) was utilized for the study, which can be coupled with different fiberoptic probes. This system can also be coupled with superhead containing ball probe or to a variety of objectives (40X, 50X and 100X) based on the type of experiment. The assembled Raman system with ball probe is shown in Figure 2.1.

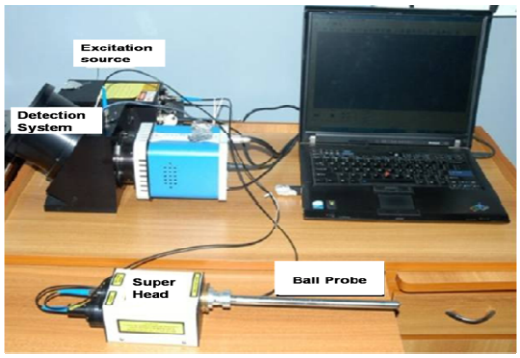


Figure 2.1: Pictorial representation of Raman system with ball probe.

In the study presented in this chapter, fiberoptic probes (Immersion and end view InPhotonics probe) and superhead with ball probe were utilized. Whereas, the Raman system coupled with superhead containing 40X objective was utilized for the *ex vivo* studies on cells; presented in the chapter 3 and 4.

The core Raman system consist of a diode laser HE 785nm (PI-ECL-785-300-FC) with 9.25" x 2.5" x 4.25" dimensions. It is made of AlGaAs, covered by thermoelectrically cooled jacket with FC connectors and 300 mW of maximum output. A high efficiency spectrograph (HE-785, HORIBA Jobin Yvon, France) with fixed 950 gr/mm grating coupled with Charged Coupled Device (CCD-1024X256-BIDD-SYN, Synapse) as a detection system. The CCD consists of 1024 X 256 pixels of 26 X 26 μm size and it is thermoelectrically cooled.

The supplementary system component called ‘Superhead’ helps in optical filtering of noise including Rayleigh light. The detection system and laser are coupled with the ‘Superhead’ using optical fibers. A ball probe can be attached to the ‘Superhead’ for remote applications especially for *in vivo* measurements. Dimension of the ball probe is 10 inches x 2.5 inches. The ball probe contains a

lens at the tip and a hollow aluminum core. As all processes, like filtering and collection of signals takes place inside the 'Superhead', the key function of the ball the probe is to deliver and collect laser light and scatter photons, respectively (Figure 2.1).

Commercially available fiberoptic probes such as an immersion and an end view probes (InPhotonics Inc., Downy St, USA) were also tested for the study. Photographic representation of the InPhotonics immersion and end view probes along with detachable sleeve is represented in Figure 2.2 and 2.3, respectively.

Figure 2.2: Pictorial representation of InPhotonics immersion probe (RPS 785/12-5)

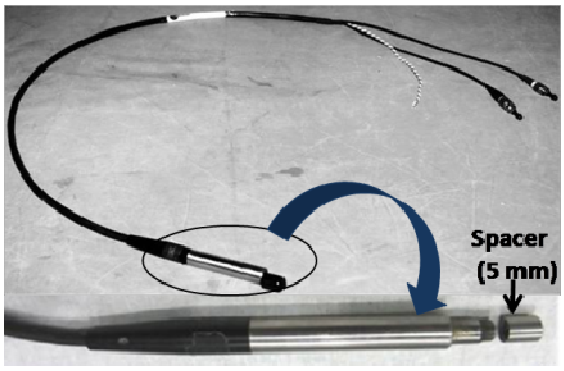


Figure 2.3: Pictorial representation of InPhotonics end view probe.

The major difference between these probes is that they have different dimensions, as well as different detachable sleeves. The dimensions of immersion and end view InPhotonics probes are 23 cm (length) x 1.5 cm (diameter) and 10

cm (length) x 1.25 cm (diameter), respectively. The InPhotonics immersion probe contains long jacket of 23 cm (length) x 1.5 cm (diameter) covering the probe (Figure 2.2). Whereas end view InPhotonics probe contains spacer (short jacket at tip) of 0.5 cm (length) x 1.5 cm (diameter) (Figure 2.3). These jackets were utilized to avoid patient to patient cross contamination.

The InPhotonics immersion probe and end view probe are coaxial and contains excitation and collection fibers. The excitation fiber is 105 µm in diameter with NA-0.40, whereas the collection fiber is 200 µm diameters and has NA-0.40. A schematic diagram of the internal optics is shown in Figure 2.4. The optical design is patent protected (U.S. Patent #5,122,127).

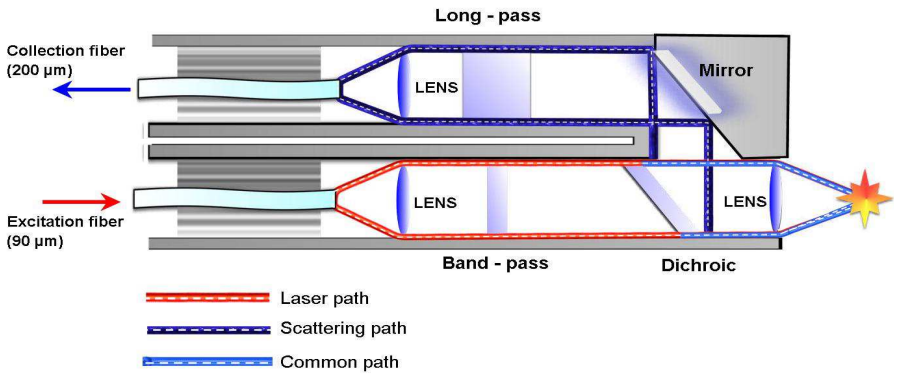


Figure 2.4: Schematic diagram of the internal optics of InPhotonics end view probe.

To optimize the overlap between the collection and excitation fibers, lenses are utilized to focus the laser light as well as scattered photons. As indicated by the manufacturer's specifications, the theoretical spot size is 105 µm and depth of field is 1 mm. Filtering and photon collection mechanisms occur in a stainless-steel sheath of 0.5 inches diameter. This makes the probe ideal for *in vivo* applications. The Figure 2.4 illustrates the beam path within the probe. A lens is

used to collimate the excitation light at the end of excitation fiber. A band-pass and long pass filter are inserted into the laser light path and collection path, respectively. While transmitting excitation light, laser noise is removed with the help of band pass filters, whereas, the long pass filter is transparent to Raman photons originating from the tissue but filters the elastic Rayleigh scattered light. Laser light is transmitted through the dichroic filter, and is eventually focused by the lens at the tip of the probe onto the sample. The backscattered tissue photons are collected by the same lens and reflected by the dichroic filter to a collection path. The long-pass filter transmits only the Stokes scattered photons and attenuates the Rayleigh signals. Finally, these Stokes photons are focused onto the collection fiber by the lens. The compact dimensions of this probe provide flexibility to the clinicians as well as comfort for patients while acquiring *in vivo* spectra.

For *ex vivo* measurements, the end view InPhotonics probe was placed on a probe holder which facilitates acquiring spectra from tissue placed on XYZ stage at different points. A pictorial representation of Raman spectroscopic setup with end view InPhotonics probe used for *ex vivo* study is shown in Figure 2.5.

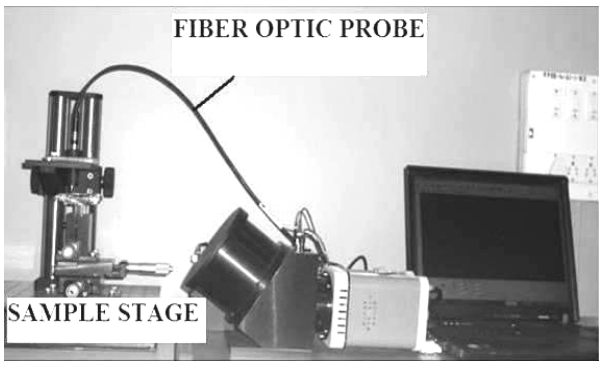


Figure 2.5: Photographic representation of Raman setup used for *ex vivo* study.

2.1.2 Sample details for *ex vivo* study

Total 27 histopathological certified cervix tissues were collected and stored in -80^{o}C until used. From these, 16 tumor tissues were collected from locally advanced cancer subjects before undergoing treatment and 11 normal tissues were collected from subject undergoing hysterectomy.

Raman spectra were measured using a fiberoptic end-view probe (InPhotonics inc., Downy St, USA) coupled HE-785 system, which is shown in Figure 2.8. Tissues were thawed and kept on a calcium fluoride (CaF_2) window. The CaF_2 window along with sample was placed on an XYZ precision stage under the illumination zone. Raman spectra were measured at different points at spacing of ~2mm. Acquisition parameters were: laser power - 50mW, acquisition time – 10 s and averaged over 5 accumulations. Experimental conditions were kept constant during all the measurements.

2.1.3 Sample details for *in vivo* study

One hundred and three (103) subjects diagnosed with gynecological cancers and planning their treatment at the Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Tata Memorial Centre, participated in the study. Out of 103, 73 and 30 subjects were cervical tumor and normal cervix cases, respectively. Raman spectra were acquired from 200 tumor cervix sites, 74 normal cervix sites, and 168 normal vaginal sites from tumor and normal subjects. Details are shown in Table 2.1. The informed and written consents were obtained from subjects prior to spectral recording. Inclusion criteria for the study consist of non-pregnant subjects having age of 30-70 and with no history of a hysterectomy.

Out of 103 subjects, 96 (93%) were post-menopausal and only 7 (~7%) were pre-menopausal subjects. Clinical details like age, family cancer history, menopausal and menstruation status were obtained through questionnaire.

Table 2.1: Total specimens details utilized in the study.

Specimen details	No. of cases	No. of spectra
Tumor cervix $	73	200
Normal Cervix #	30	74
Normal vagina of tumor subjects $	51	104
Normal vagina of normal subjects #	25	64
Total	103	442
Note: $ and # marked cases consist of 51 and 25 common cases, respectively.		

Speculum was inserted in the vagina so as to observe a cervix which was cleaned with saline solution.

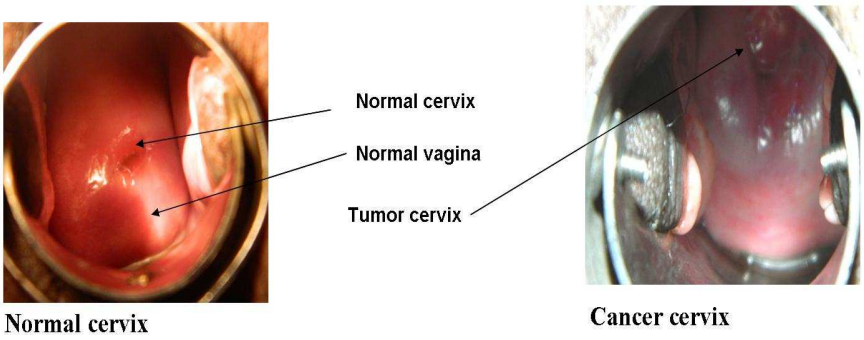


Figure 2.6: Pictorial representation of in vivo sites (normal, tumor and vagina cervix)

'White light imaging' is one of the ways to conduct *in vivo* cervical cancer Raman spectroscopic study where colposcope is essential, while another approach is utilized in this study. In this approach, the *in vivo* Raman spectra were recorded from the different o'clock position of cervix, for eg: 12'oclock, 3'oclock etc. Recorded cites were histopathologically further verified. Moreover, *in vivo* study

was carried out on full grown tumor cervix, which can be easily seen and not on pre-cancer patches, where white light imaging is essential. These all procedures were done under clinical supervision. Multiple *in vivo* Raman spectra (3-6) were recorded from gross tumor, normal cervix and vagina of normal and tumor subjects under clinical supervision. Raman spectra were acquired by placing probe perpendicular to the surface of site. To avoid contamination in the subjects, prior to spectral recording, the probe was disinfected with CIDEX (Johnson and Johnson, Mumbai, India) solution and wrapped in parafilm. Photographic representation of normal cervix, normal vagina and tumor cervix is shown in Figure 2.6.

2.1.4 Standardization of Raman spectral acquisition

Initially Raman system with superhead attached ball probe was utilized for Raman spectral acquisitions. Pictorial representation of Raman system with ball probe is shown in Figure 2.1. The superhead with ball probe were heavy, extremely inconvenient and impractical for *in vivo* application. Furthermore, Raman spectra measured by this probe were contaminated by high background and noise from the ball probe, interfering with the fingerprint region of the spectrum. A raw spectrum acquired by the ball probe at 50 mW, 10 s acquisition time, over 5 cycles is shown in Figure 2.7.

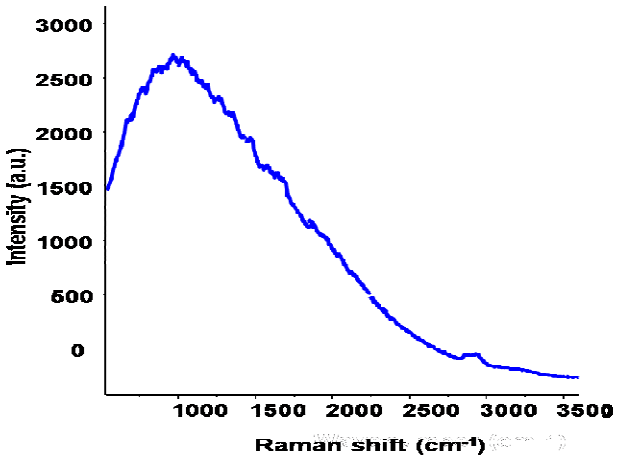


Figure 2.7: Raw *in vivo* spectrum of tissue acquired by ball probe

Due to the high noise, poor spectral quality and patient's discomfort, the study using ball probe was discontinued.

The utility test of the InPhotonics immersion probe for spectral acquisition from tissue specimens was carried out. The pictorial representation of InPhotonics immersion probe is shown in Figure 2.2. The raw spectrum measured with the InPhotonics immersion probe is shown in Figure 2.8.

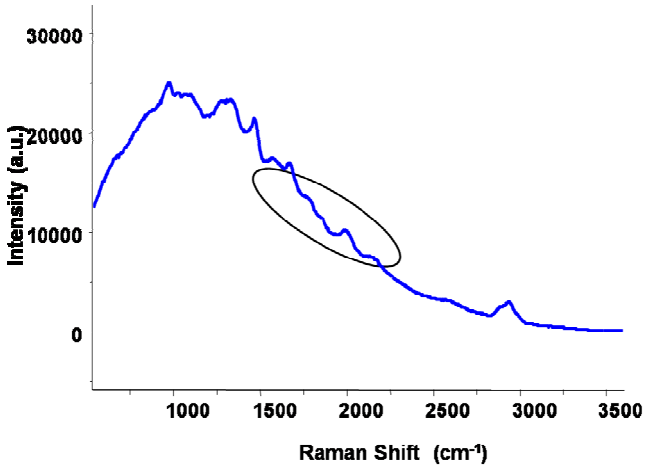


Figure 2.8: Raw spectrum of tissue acquired by InPhotonics immersion probe.

Several additional bands, probably originating from the fiber optic probe were observed. Since, the InPhotonics immersion probe gave poor quality spectra; this probe was abandoned. As an alternative, the applicability of an end view Raman probe was explored for *in vivo* cervical cancer study. This probe has already being successfully utilized in our lab for an oral cancer program.

Before employing the probe for *in vivo* study, *ex vivo* spectra were acquired from cervix tissues to verify the reproducibility of spectra. The photographic representation of Raman setup used for *ex vivo* study is shown in Figure 2.5. A typical *ex vivo* raw spectrum of cervix tissue is shown in Figure 2.9. A good quality *ex vivo* tissue spectrum could be acquired by the end view probe; hence this probe was procured for *in vivo* cervical cancer study.

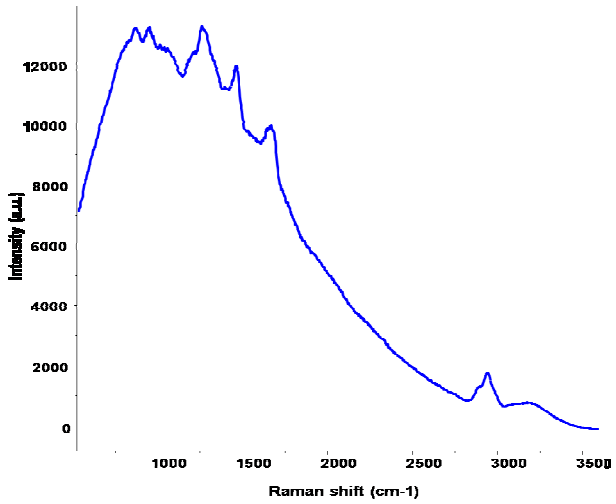


Figure 2.9: Raw spectrum of cervix tissue acquired by InPhotonics end view probe.

After standardization and verifying the reproducibility of Raman spectra, the end-view probe was utilized for *in vivo* cervical cancer study. But, it was

observed that, since the end view probe is 5 inches in length, it had poor grip on probe while recording *in vivo* spectra from the cervix sites.

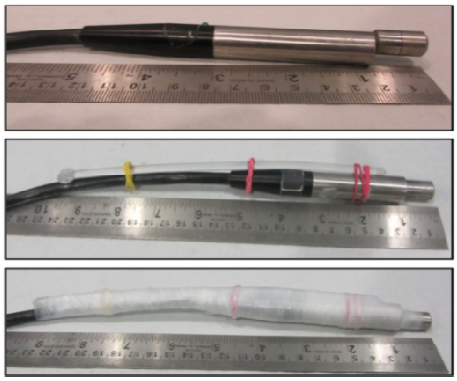


Figure 2.10: Pictorial representation of end view probe's modification for *in vivo* study

To resolve this problem, a rod was tied to the probe with rubber bands so as to increase the length to 9 inches. To avoid patient to patient cross infection, this probe was further wrapped with parafilm (Figure 2.10). A typical raw *in vivo* spectrum of cervix sites, acquired by the end view probe is shown in Figure 2.11.

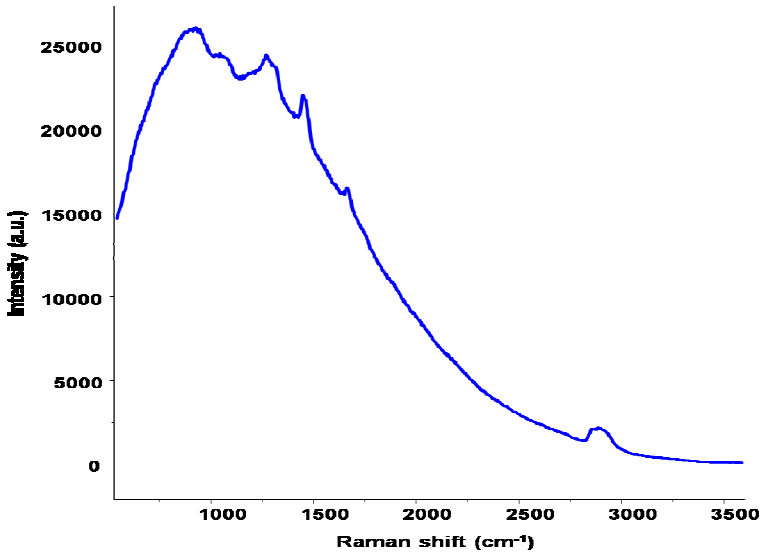


Figure 2.11: Raw *in vivo* spectrum of cervix acquired by InPhotonics end view probe

2.1.5 Raman spectral pre-processing

Raman spectra were pre-processed as per standard protocol. Typical *in vivo* spectra at different pre-processing steps are shown in Figure 2.12. The raw Raman spectra, acquired under 785-nm laser excitation, are composed of Raman signal, autofluoroscence background and noise. Raw spectra were pre-processed by utilizing Labspec 5.0 software (HORIBA Jobin Yvon). Initially, all spectra were corrected for the wavelength-dependent intensity response of the system using a calibration standard (standard reference material number- 2241; NIST, Gaithersburg, MD, USA). This was achieved by measuring the calibration standard spectrum which was further divided with raw white light spectrum so as to remove the signals associated with the instrument response [127].

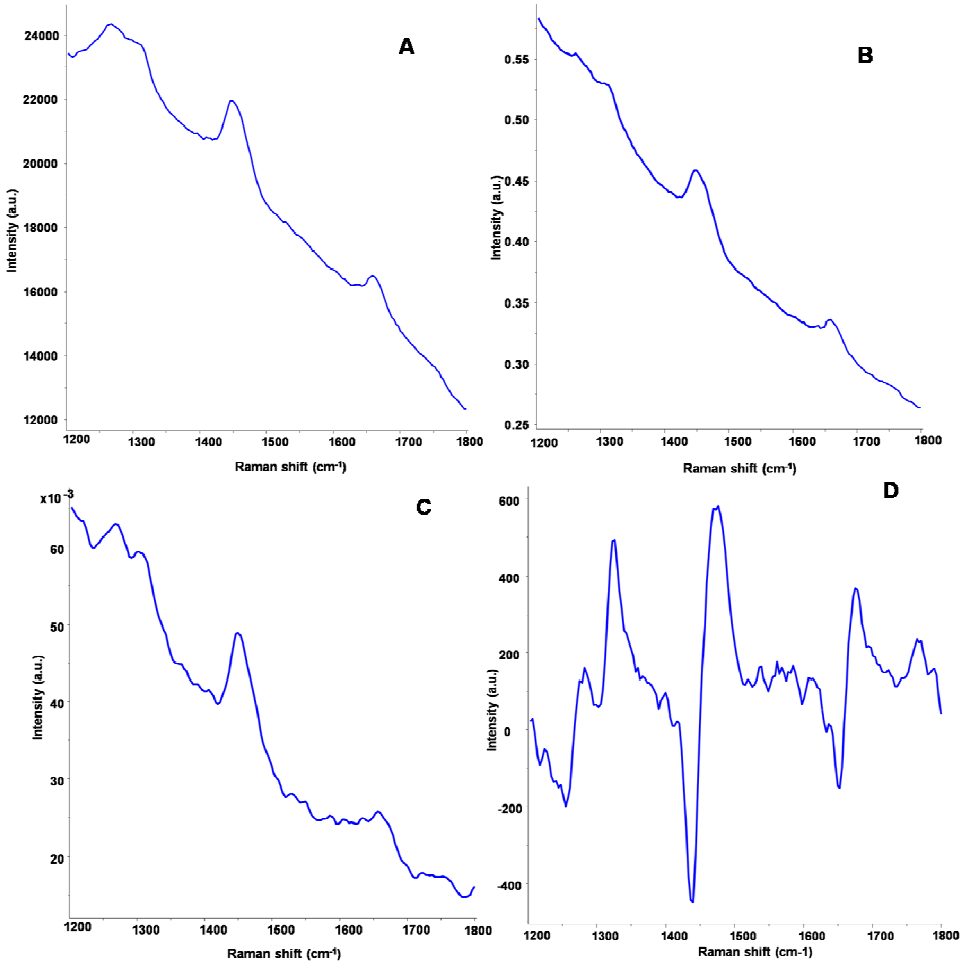


Figure 2.12: Representative *in vivo* spectrum at different pre-processing steps (A) Raw spectrum (B) CCD response corrected (C) Background corrected (D) First Derivative

The spectral contributions from the background i.e (optical elements, air, etc) were obtained by acquiring air spectra under identical experimental conditions. The response corrected background spectrum was subtracted from the response corrected raw spectrum. Slow moving autofluorescence background was removed by computing the first derivative spectrum using Savitzky–Golay filter mechanism (window size-3) [128, 129]. The prime objective of the first derivative correction was to construct a spectral profile depicting point by point variation of

spectral intensity over a moving window of 3 points. This transformation provides peak profile in the spectra irrespective of optical response related intensity. Background corrected spectra were interpolated and 1st derivatized, which was followed by vector-normalization. First derivatized, vector-normalized spectra were then subjected to multivariate statistical tool PC-LDA.

2.1.6 Computing *ex vivo* and *in vivo* average spectra

Average spectra were computed by averaging variations on Y-axis, keeping the X-axis constant using background subtracted spectra (prior to derivatisation) for each class. This was carried out using baseline correction by fitting a 5th order polynomial function. Spectral comparisons across all groups were done by using these baseline corrected average spectra. The difference spectrum was calculated from vector normalized baselined spectra for comparisons across different groups.

2.1.7 Multivariate statistical analysis

Data was analyzed by Principal Component-Linear Discriminant Analysis (PC-LDA). Details of PC-LDA have been elaborated in chapter 1 under the section 1.4 of data analysis. Briefly, PCA aims to identify features that represent variance in the data; LDA provides data classification based on an optimized criterion which is objective for more class separability. LDA transformation matrices are generated and it further identifies eigenvector or LDA components of this classification criterion. The scree plots depict the variance (% correct classifications) accounted for the total number of LDA components selected for analysis. The outcomes of PC-LDA are represented in the form of a confusion matrix, where diagonal elements are true positive and non-diagonal elements are

false positive predictions. The confusion matrix aids to understand the separation within the groups which are acquired by accounting for the contribution of all selected factors used for analysis. PC-LDA results are also represented in the form of scatter plots, generated by plotting various combinations of scores of factors. The performance PC-LDA diagnostic models were further validated in an unbiased method by leave-one-spectrum-out, cross-validation (LOOCV). In LOOCV methodology, one spectrum is held out from the data set and remaining data is used to redevelop PC-LDA model. Test prediction was also used to validate the models. This process is repeated until all withheld spectra are classified. Algorithms for these analyses were implemented in MATLAB (Mathworks Inc.) based in-house software [101]. Different spectral regions like full-range, fingerprint and high-wave-number were explored for classification. The best classification was achieved using the 1200-1800 cm^{-1} region, and as this region is least influenced by fiber interferences, it was therefore selected for analysis. Since previous studies have demonstrated the efficacy of the 1200-1800 cm^{-1} region in classifying normal and malignant oral cancers and as it is less influenced by fiber signals, same region for further analyzed was employed [127, 130].

2.2 Results and discussion

2.2.1 The *ex vivo* cervical cancer study

Previous studies have demonstrated that spectra of normal tissues are dominated by collagenous type proteins, while tumor tissues are rich in non-collagenous protein and nucleic acids [106, 120]. This study was undertaken to evaluate the

reproducibility of the spectral features of normal and tumor tissues. Using a fiberoptic probe coupled Raman spectroscope, spectra of normal and cancer cervical biopsies were acquired and analyzed.

2.2.1.1 Spectral profiles of normal and tumor cervix tissue

Vector normalized average *ex vivo* spectra of normal and tumor tissues along with their standard deviations are shown in Figure 2.13.

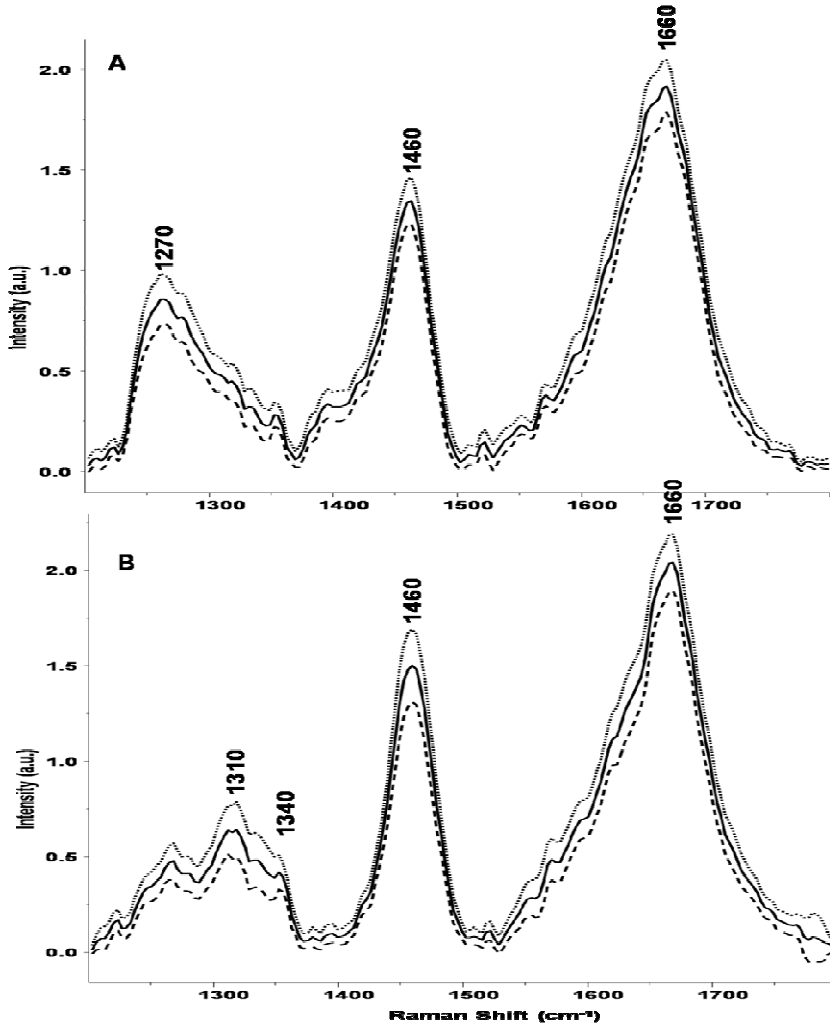


Figure 2.13: (A) Average *ex vivo* spectra of normal and (B) tumor cervix tissue (Solid line mean spectra, dotted line- mean + standard deviation, broken line- mean – standard deviations)

Solid, dotted and broken lines represent mean, mean + standard deviation and mean - standard deviation, respectively. Minor intensity variations were observed within the group. Spectral characteristics of collagen like features, amide III with strong and broad amide I, were observed in the normal mean spectrum. The tumor spectrum exhibited a strong and sharp amide I, a minor shift in δ CH_2, 1460 cm^{-1} and a band at 1340 cm^{-1}, indicative of DNA and non-collagenous proteins. These findings corroborates with previous studies, demonstrating the reproducibility of spectra [106, 120, 131].

To illustrate the spectral differences between normal and tumor; a difference spectrum was computed by subtracting the average spectrum of normal from tumor. The difference spectrum is shown in Figure 2.14. The positive peaks of difference spectrum are from the average tumor spectrum and negative bands are due to the mean normal spectrum. Positive peaks of proteins like amide I (1660 cm^{-1}), δCH_2 deformation (1450 cm^{-1}), and DNA (1340 cm^{-1}) can be seen in mean tumor cervix spectrum whereas negative peak of protein (1280 cm^{-1}) signify a higher collagenous protein presence in normal cervix tissue.

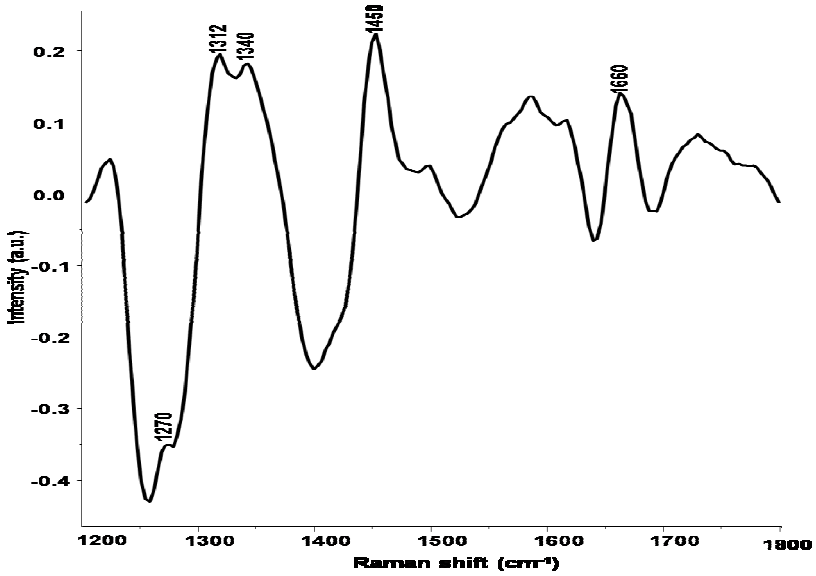


Figure 2.14: Difference spectra of tumor-normal cervix tissue

2.2.1.2 Classification of normal and tumor cervix tissue spectra

To evaluate the feasibility of classification of tumor and normal spectra, the first derivatives of pre-processed spectra were subjected to supervised PC-LDA followed by leave-one-out cross-validation (LOOCV). The scree plot represents the variance or percent correct classifications accounted for the total number of factors selected for analysis. Four factors, contributing ~95 % of classification, were used for analysis, as shown in Figure 2.15 A.

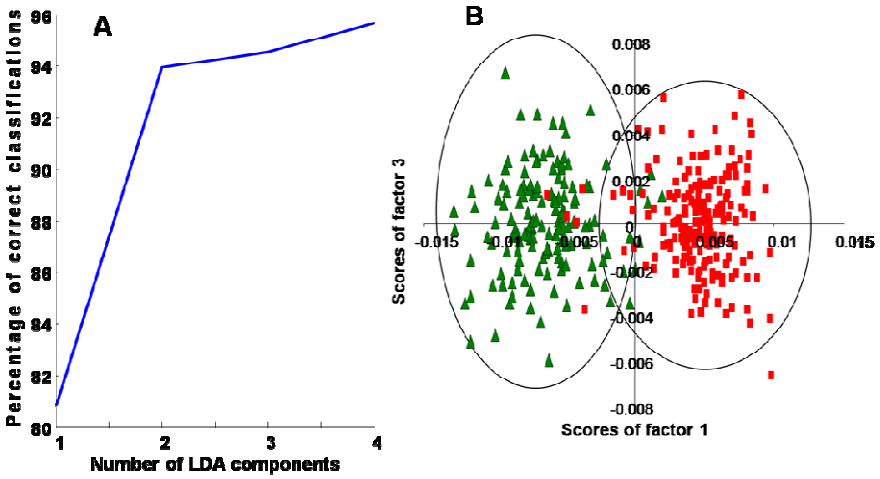


Figure 2.15: PC-LDA of normal and tumor (A) Scree plot (B) Scatter plot (tumor cervix (■) and normal cervix (▲).

The scatter plot is shown in Figure 2.15 B. It illustrates two distinct clusters belonging to normal and tumor cervix spectra.

Table 2.2: Principal Component-Linear Discriminant Analysis and leave-one-out cross validation of *ex vivo* normal (N) and tumor (T) cervix tissue (Diagonal elements are true positive predictions and ex-diagonal elements are false positive predictions).

Standard Model			
	Normal	Tumor	% efficiency
Normal	142	6	96
Tumor	13	188	94
Leave-One-Out Cross-Validation			
	Normal	Tumor	% efficiency
Normal	139	9	94
Tumor	18	183	91

PC-LDA results are also summarized in Table 2.2. As can be seen, 142/148 normal and 188/201 tumor spectra were correctly classified. Six of the

tumor spectra and 13 of normal spectra were misclassified as normal and tumor, respectively. LOOCV was also carried-out to evaluate the classification efficiency of the model. As mentioned earlier, cross-validation is a method for assessing reliability of a predictive model with a hypothetical validation set, leave-one-out (LOO). 139/148 and 183/201 were correctly classified as normal and tumor, respectively. Only 9 spectra out of 148 normal were misclassified and 18 tumor spectra were misclassified. An average classification efficiency of 92.5 % was observed.

2.2.2 The *in vivo* cervical cancer study

This study was carried out in two parts, in the first part the classification of normal and tumor *in vivo* spectra was explored whereas in second part the utility of vaginal sites as an internal control was tested.

2.2.2.1 Spectral profiles of normal and tumor cervix sites

Vector normalized average *in vivo* spectra of tumor cervix (T), normal cervix (N), vaginal sites of normal cervix (NVN) and tumor cervix subjects (NVT) are illustrated in Figure 2.16. Mean spectra of normal cervix and vagina exhibit characteristic spectral features of amide III and strong and broad amide I. These can be attributed to collagenous proteins. Prominent features of tumor, with respect to normal spectrum, are strong and sharper amide I, a minor shift in δ CH_2 (1460 cm^{-1}) and a distinct band at 1340 cm^{-1} which are indicative DNA and non-collagenous proteins. These findings corroborate earlier *ex vivo* and *in vivo* Raman spectroscopic studies on cervical cancers [106, 107, 120, 131]. Differences in the form of shifts and intensity related variations were observed.

To highlight the spectral differences different groups, difference spectra were computed (Figure 2.17). To explore the spectral differences between tumor cervix (T) and normal cervix spectra (N), difference spectra was computed by subtracting the mean spectrum of normal from the mean tumor spectrum (Figure 2.18 A). The positive peaks in the difference spectrum are from the mean tumor cervix spectrum and negative bands are due to mean normal cervix spectrum. Strong positive peaks of protein like amide I (1662 cm^{-1}), δ CH_2 deformation (1450 cm^{-1}), and DNA (1340 cm^{-1}) can be seen in the mean tumor cervix spectrum whereas the negative peaks of protein (1280 cm^{-1}) signify a higher collagenous protein presence in normal cervix. The difference spectra of T-NVT, T-NVN and NVT-NVN are illustrated in Figure 2.17 B, C and D, respectively. It was observed that difference spectra of T-NVT and T-NVN showed a similar kind of spectral profile to that of T-N. Observed positive peaks at 1660 cm^{-1}, 1450 cm^{-1} and 1340 $cm^{-1}$of difference spectra were characteristic of tumor cervix indicated increased DNA and protein while negative peaks at 1280 cm^{-1} and 1240 cm^{-1} indicate collagenous protein. The difference spectra of NVT-NVN showed minor variation in amide I and δCH_2 peaks .

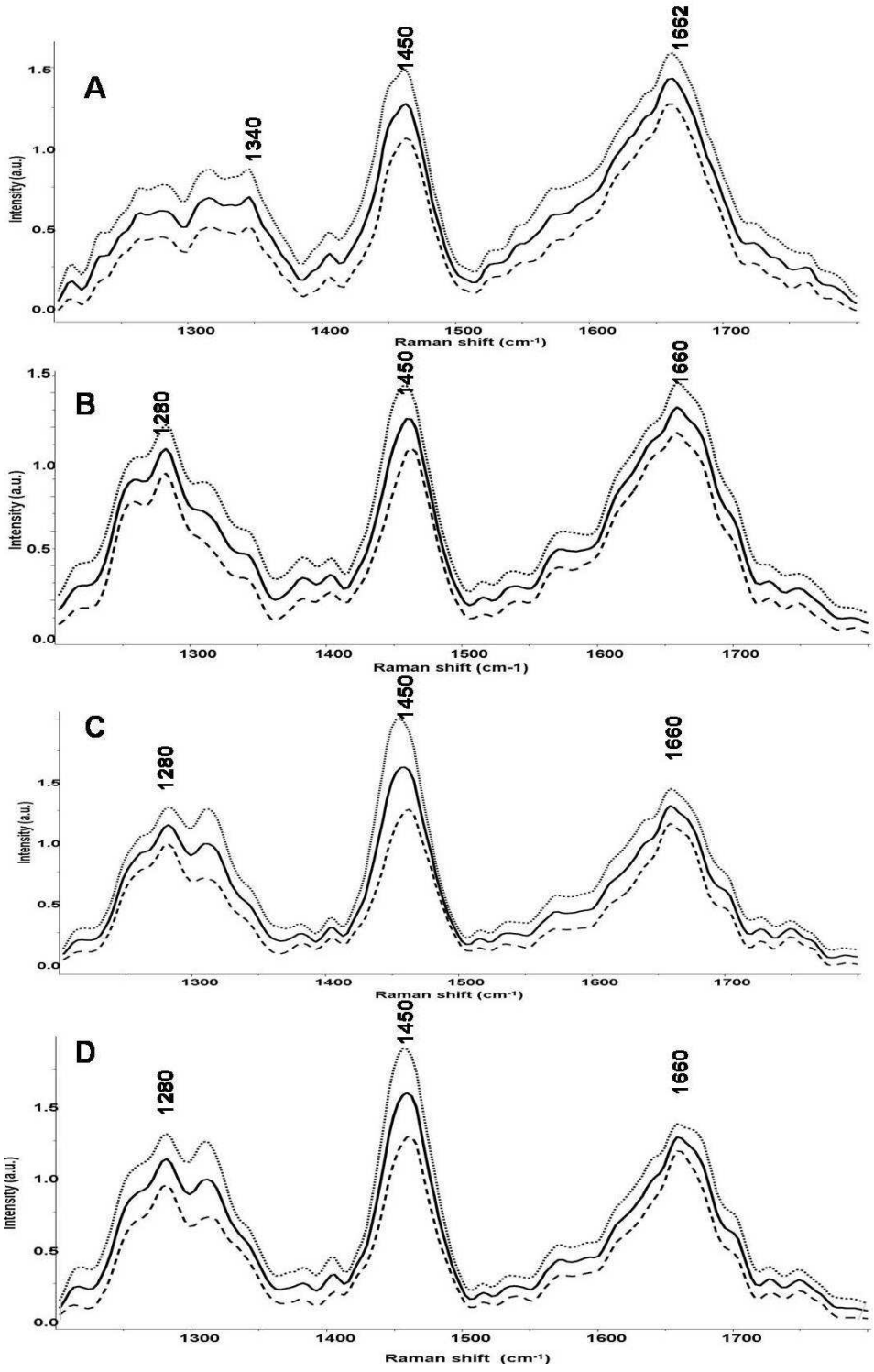


Figure 2.16: ***In vivo*** **mean Raman spectra of (A) cervical tumor (T), (B) normal cervix (N), (C) vagina of normal cervix subjects (NVN) and (D) vagina of cervical tumor subjects (NVT).**

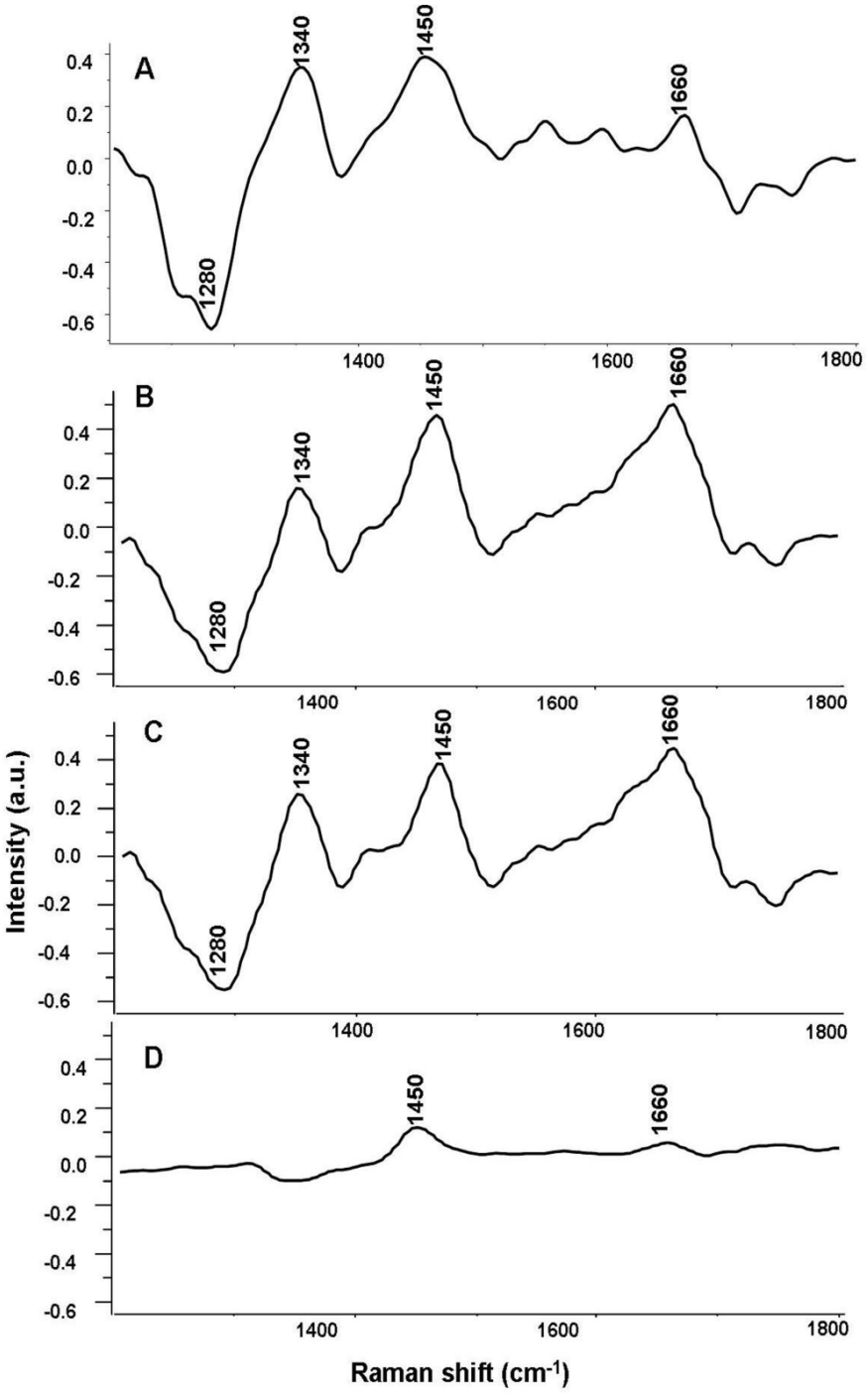


Figure 2.17: Difference spectra [A] Cervical tumor-normal cervix [B] Cervical tumor - vagina of cervical tumor, [C] Cervical tumor- vagina of normal cervix, [D] Vagina of cervical tumor- vagina of normal cervix.

2.2.2.2 Classification of normal and cancer sites

To explore the classification between normal and cervical cancer, 61 subjects (154 spectra) were enrolled in the study. Out of 154, 80 spectra were acquired from cervical tumor (T) of 31 cervical cancer subjects and 74 spectra from uninvolved normal cervix (N) of 30 subjects (Table 2.3).

Table 2.3: Sample utilized for classification of normal (N) and cervical tumor (T).

Sample details	Number of subjects	Number of spectra
Normal cervix & vagina (N)	31	74
Tumor (T)	30	80

To determine the feasibility of classification of normal and cervical tumor, the first derivative preprocessed spectra were subjected to PC-LDA followed by leave-one-out cross-validation (LOOCV).

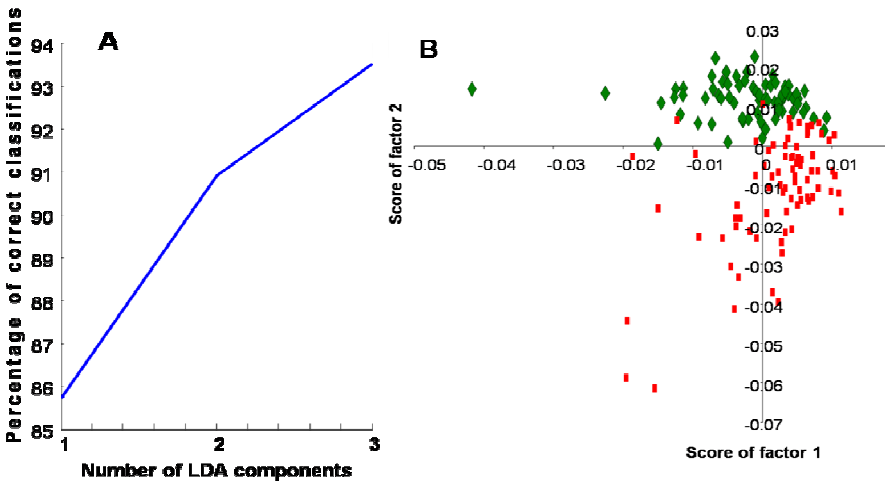


Figure 2.18: Classification of tumor and normal cervix (A) Scree plot. (B) Scatter plot (cervical tumor (■) and normal cervix (◆).

The scree plot depicts the variance/percent correct classifications; accounting for the total number of factors selected for analysis and is shown in

Figure 2.18 A. Three factors, contributing ~93.5 % of classification were used for analysis. The scatter plot shown in Figure 2.18 B depicts exclusive clusters corresponding to normal and cervical tumor.

These results have also been summarized in Table 2.4. As can be seen, 70/74 normal and 80/80 tumor spectra were correctly classified. None of the tumor spectra were misclassified, whereas 4 normal spectra were misclassified as tumors. LOOCV was also executed to evaluate classification efficiency of the model and are shown in Table 2.4 B. Only 5 spectra out of 74 normal cervix sites were misclassified and all tumor spectra were correctly classified. An average classification efficiency of 96.5% was observed.

Table 2.4: PCLDA of normal cervix and cervical tumor (A) Standard model, and (B) leave-one-out cross validation (Diagonal elements are true positive predictions and ex-diagonal elements are false positive predictions).

Standard Model		
	Normal (N)	Tumor (T)
Normal (N)	70	4
Tumor (T)	0	80
Leave-One-Out Cross-Validation		
	Normal (N)	Tumor (T)
Normal (N)	69	5
Tumor (T)	0	80

2.2.2.3 Classification among all controls (normal cervix, normal vaginal sites of normal and cancer subjects)

As mentioned earlier, multiple variables linked with cervical cancer like HPV infections, hormonal status, menopause, race, ethnicity, body mass index, parity, socio economical status have been explored, using both high wavenumber and fingerprint regions. Further careful validations on diverse population and larger

cohort are required for translation of this technology into clinics. Cervical cancer subjects in developing countries like India are very often presented at advanced stages (stage IIA and above) [132] and in such cases, as majority of cervix is diseased, contains no normal cervix site to acquire control spectra. Therefore, healthy cervix of subjects having other gynecologic cancers (uterine or ovarian) undergoing hysterectomies are used as controls, in which subject accrual is often a major constrain. Histological similarities of vagina and cervix (ectocervix) are also known [131]; moreover, malignancy-associated changes/cancer field effects are not reported in cervical cancers. Therefore, as an alternative vagina can be utilized as an internal control, especially in screening camps where colposcopy may not be available.

To explore the utility of the vagina as an internal control, 230 spectra from 66 subjects were utilized. Among 230 spectra, 74 spectra were from uninvolved normal cervix sites of 30 gynecological cancer subjects (N), 64 spectra were acquired from uninvolved vaginal sites of 20 normal cervix of other gynecological cancer subjects (NVN) and 92 spectra were from 36 cervical cancer subjects (NVT) (Table 2.5).

Table 2.5: Samples utilized for classification of all internal controls.

Sample details	**Number of subjects**	**Number of spectra**
Normal Cervix (N)#	30	74
Normal vagina of normal cervix case (NVN)#	20	64
Normal vagina of tumor cervix case (NVT)	36	92
Total number of cases	66	230
Note: # marked contain 20 common cases		

To explore the variations between the control group first derivative pre-processed spectra were used. 74 spectra from normal cervix (N), 64 from vagina of normal subjects (NVN) and 92 from vagina of tumor subjects (NVT) were analyzed by PC-LDA using 7 factors. The 7 factors contributed to only 54% of classification shown in the scree plot (Figure 2.19 A).

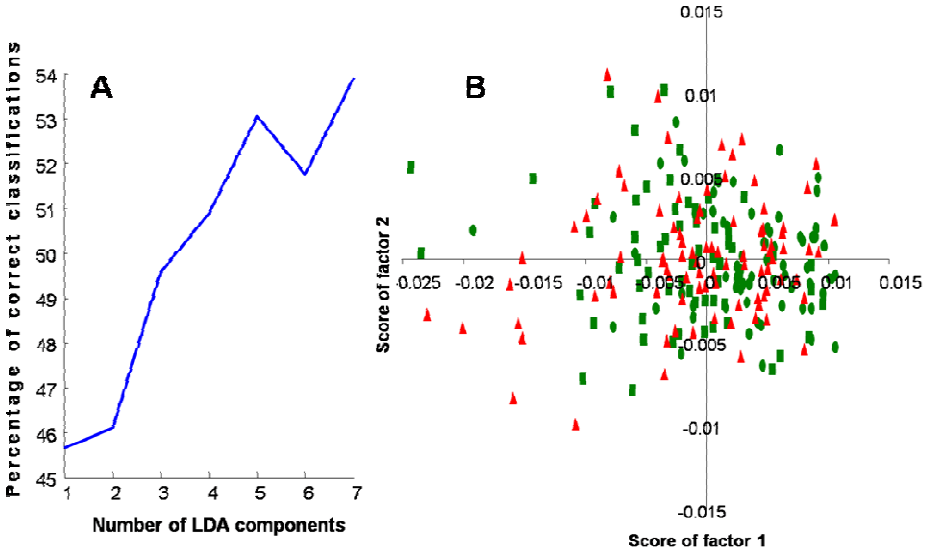


Figure 2.19: Exploring internal control- PCLDA of normal cervix, vagina of tumor cervix and vagina of normal cervix (A) Scree plot, (B) Scatter plot (normal cervix (●), vagina of normal cervix subjects (■) and vagina of normal cervix subjects (▲).

The scatter plot is shown in Figure 2.19 B, which exhibits huge misclassification among normal cervix (N), vagina of normal cervix subjects (NVN) and vagina of cervical tumor subjects (NVT).

The findings of PC-LDA are also shown in the confusion matrix for the standard model and LOOCV in Table 2.7 A and B, respectively. In the case of the standard model 49/74 normal cervix spectra (N), 41/64 vagina spectral sites of normal subjects (NVN) and 34/92 vagina spectral sites of tumor subjects (NVT)

were correctly classified. In the case of LOOCV, 44/74 normal cervix spectra (N), 25/64 vagina of normal subjects (NVN) and 21/92 vagina of tumor subjects (NVT) were correctly classified. Thirteen and 17 of normal cervix (N) were misclassified with vagina of normal subjects (NVN) and vagina of tumor subjects (NVT), respectively and in the case of vagina of normal subjects (NVN), 9 and 30 spectra were misclassified with normal cervix (N) and vagina of tumor subjects (VT), respectively. For vaginal sites of tumor subjects (NVT), 36 spectra were misclassified with normal cervix (N) and 35 with vagina of normal subjects (NVN). The higher misclassification was observed between the spectra of normal cervix (N), vagina of normal subjects (NVN) and vaginal sites of tumor subjects (NVT) is indicative of the biochemical similarities of these groups. The findings suggest that vagina can be used as internal control on similar lines to oral and breast cancers wherein, contralateral and uninvolved areas are employed as controls, respectively. This approach may also help to circumvent possible influence of hormonal status, menopausal status, age and parity. Also spectral acquisition does not require colposcope at the site.

As the spectra of vagina of normal subjects (NVN) and vagina of tumor subjects (NVT) show similarity, these spectra were grouped together and referred to them vagina spectra (V) in our subsequent evaluation of vagina as control. It is also important to note that among 93 subjects, 87 (93%) cases were post-menopausal and only 6 (7%) were pre-menopausal. Hence, menopausal status may not have influence on current results.

Table 2.6: Principal Component-Linear Discriminant Analysis and leave-one-out cross validation of normal cervix (N), vagina of normal cervix subjects (VN) and vagina of tumor cervix subjects (VT).

Standard Model			
	Normal cervix (N)	Vagina of normal cervix subjects (NVN)	Vagina of tumor cervix subjects (NVT)
Normal cervix (N)	49	13	12
Vagina of normal cervix subjects (NVN)	9	41	14
Vagina of tumor cervix subjects (NVT)	32	26	34
Leave-one-out cross-validation			
	Normal cervix (N)	Vagina of normal cervix subjects (NVN)	Vagina of tumor cervix subjects (NVT)
Normal cervix (N)	44	13	17
Vagina of normal cervix subjects (NVN)	9	25	30
Vagina of tumor cervix subjects (NVT)	36	35	21

2.2.2.4 Classification of tumor cervix, normal cervix and normal vaginal sites-the in vivo study

To evaluate the utility of vagina as an internal control, a total of 442 spectra were utilized. Out of 442, 200 tumor spectra were acquired from 73 cervical cancer subjects (T), 74 normal cervix spectra from uninvolved normal cervix sites of 30 gynecological cancer subjects (N), 168 vaginal spectra were acquired from uninvolved vaginal sites of 76 subjects (V) (Table 2.6).

In order to evaluate the efficacy of Raman spectroscopic methods in discriminating tumor conditions against control groups, spectra of tumor cervix (T), normal cervix (N), vagina (V) were subjected to PC-LDA.

Table 2.7: Sample details for classification to evaluate internal controls.

Sample details	Number of subjects	Number of spectra
Tumor (T)$	73	200
Normal Cervix (N)	30	74
Vagina (V)$	76	168
Total number of cases	103	442
Note: $ marked contain 73 common cases		

In the first step, 28 spectra from tumor (T), 34 from normal cervix (N), 24 spectra of vaginal (V) were employed to build a standard model by utilizing seven factors. The scree plot exhibited a total 5 factors contributing to 85% of correct classification as shown in Figure 2.20A. The scatter plot, shown in Figure 2.20B, illustrated two clusters belonging to tumor and control spectra (normal cervix and vagina).

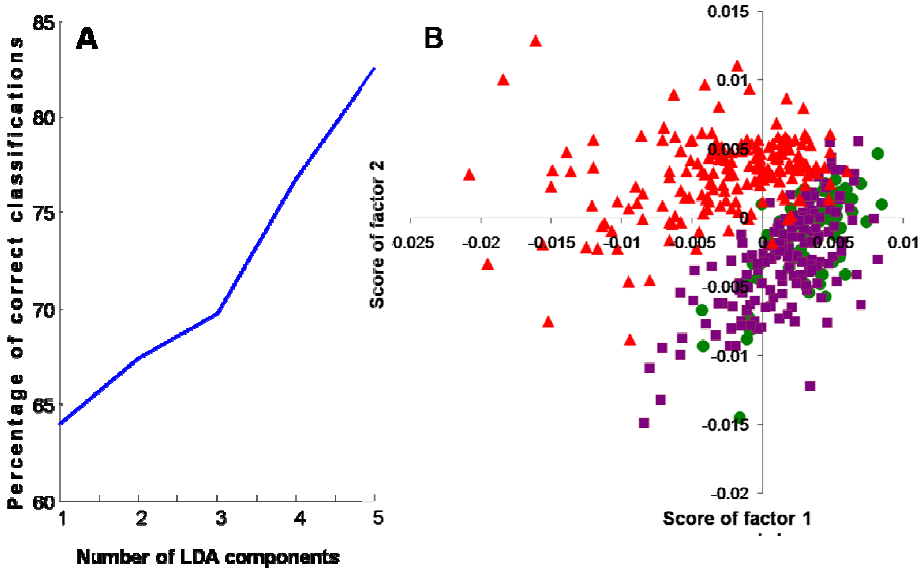


Figure 2.20: Verifying internal control- PCLDA of cervical tumor, normal cervix, vagina (A) Scree plot, and (B) Scatter plot (cervical tumor (▲), normal cervix (●) and vagina (■).

Clear classification was observed among the clusters belonging to tumor (T) and controls i.e. normal cervix sites (N), vagina (V). Spectra from normal cervix sites (N) and vagina (V) exhibited very high overlap. Findings of PC-LDA are also shown in confusion matrix of Table 2.8 A and B. As can be seen, 25/28 cervical tumor, 27/34 normal cervix and 19/24 vagina spectra were correctly classified. In the case of LOOCV, 24/28 cervical tumor (T), 22/34 normal cervix (N) and 8/24 vaginal spectra (V) were correctly classified. 12/34 spectra of normal cervix were misclassified with vagina (V) and 15/24 vaginal (V) exhibited the misclassifications with normal cervix (N) spectra.

Table 2.8: Verification of internal control-Principal Component-Linear Discriminant Analysis, leave-one-out cross validation and test prediction of tumor cervix (T), normal cervix (N) and vagina (V).

A. Standard Model			
	Tumor (T)	Normal (N)	Vagina (V)
Tumor (T)	25	2	1
Normal (N)	0	27	7
Vagina (V)	0	5	19
B. Leave- one-out cross-validation			
	Tumor (T)	Normal (N)	Vagina (V)
Tumor (T)	24	4	0
Normal (N)	0	22	12
Vagina (V)	1	15	8
Test prediction			
	Tumor (T)	Normal (N)	Vagina (V)
Tumor (T)	170	5	3
Normal (N)	1	30	18
Vagina (V)	0	53	90

The predictive efficiency of the standard model was evaluated by using 178 tumors (T), 49 normal cervix (N), and 143 vaginal sites (V) as independent test data set. In this case, 170/178 tumor (T), 30/49 normal cervix (N) and 90/143 vagina spectra (V) were correctly predicted. However, 18/49 normal cervix (N)

were misclassified as vagina (V) and 53/143 vagina (V) were misclassified as normal cervix (NC) (Table 2.8 C). Higher misclassifications between normal cervix (N) and vagina (V) once again suggest similarities between the biochemical compositions. This further supports the applicability of vagina as internal control. Out of 178 tumors, 5 and 3 were misclassified as normal cervix (N) and vaginal sites (V), respectively. The observed minor misclassifications of tumor (T) as normal cervix (N) may be attributed to heterogeneity of tumors. As spectra were recorded at several points and few of the sites may be from islands of normal in tumors.

In this study, the efficacy of Raman spectroscopic classification of normal and cervical cancers in Indian population was evaluated and the utility of vaginal sites as an internal control was explored. The PC-LDA of normal (N) and tumor (T) sites gave classification efficiency of 96.5% was observed. On the other hand, PC-LDA of normal cervix (N), and vagina of tumor subjects (VT) and vagina of normal subjects (VN) showed higher misclassifications, suggesting similarities in biochemical composition among controls. PC-LDA of tumor (T), normal cervix (N), and vagina (V) showed classification between tumors and all controls i.e normal cervix (N) and vagina (V). Large misclassifications between the control spectra were observed. This further supports the utility of vagina as an internal control. The study also demonstrates that Raman spectroscopy may be used for improving cervical cancer diagnosis by incorporating an internal control like vagina to circumvent the influence of parameters like hormonal status,

menopausal status, and age, besides requirement of colposcope especially for mass screening in make shift camps.

2.3 Summary

The first section of the chapter discussed the standardization of *in vivo* Raman spectral acquisitions utilizing *ex vivo* cervix tissues, while the second section elaborated the utility of a fiberoptic probe coupled Raman spectroscope for *in vivo* application in cervical cancers. To the best of our knowledge, for the first time, the feasibility of acquiring good quality *in vivo* cervix Raman spectra from the Indian population was demonstrated. The utility of vagina as an internal control have also been explored.

The work presented in this chapter has been summarized as follows:

- ✓ Raman spectroscope coupled with a fiberoptic probe was procured and assembled in the laboratory. Data acquisition and analysis protocol was standardized by utilizing *ex vivo* measurements on normal and tumor cervix tissues.

- ✓ Spectral reproducibility has been established. Normal cervix tissues spectra were rich in collagenous type of proteins while tumor tissue spectra were predominated by non-collagenous type of proteins and nucleic acid. Standard models were developed and evaluated with leave-one-out-cross-validation. It was observed that standard models of normal and tumor tissue spectra exhibited the predictions efficiency of 94 and 91 %, respectively.

- ✓ The feasibility of objectively classifying tumor and normal cervix *in vivo* sites was tested. Prediction efficiency of 94.5 and 100 % for normal and tumor *in vivo* sites was observed, respectively. Finding suggests that applicability of Raman spectroscopic methods for objective, noninvasive and rapid cervical cancers diagnosis and corroborates earlier reports.

- ✓ Utility of normal vaginal sites as an internal control was also explored. Multivariate statistical analysis of normal cervix and vaginal sites of tumor and normal subjects was carried out. The findings suggest that vaginal sites can be used as internal control, where the normal cervix sites may be unavailable due to advancement of disease.

3 RAMAN SPECTROSCOPY OF EXFOLIATED CERVICAL CELLS SPECIMENS

Introduction

As mentioned in chapter 1 in section 2.3, Fourier Transform Infrared (FTIR) spectroscopic studies have demonstrated differences between normal and cancerous exfoliated cell specimens [59, 61-63, 67, 68]. But the presence of water in biological specimens is a serious hurdle in FTIR spectroscopy [133]. In order to overcome this problem, sample drying has been practiced. However, it is well known that sample drying may alter the morphology and biochemical composition of cells [134, 135]. The vibrational signals of such samples might not represent the true biochemical state of the cells. Another limitation of specimen drying, after spectroscopy, the cell specimens cannot be used for Pap staining. Therefore, in such cases, parallel sampling has been carried out. But, parallel sampling for spectroscopy and staining may not be ideal for cytological correlation, since the abnormal cell content in an 'abnormal' smear collected from the same patient may vary.

In contrast, Raman spectra are minimally influenced by water and hence Raman spectroscopy require minimal or no sample preparations. Therefore, a single specimen can be used for both spectroscopy and Pap staining, facilitating better cytological correlation. Thus, Raman spectroscopy is better suited for exfoliated cell study. In this chapter, a Raman spectroscopic approach to differentiate exfoliated cervical cell specimens and the influence of diverse factors on the classification was explored. This chapter is divided into three sections: the first section describes Raman spectroscopic classification of untreated normal and abnormal specimens. The second section of the chapter is devoted to Raman

studies of Red Blood Corpuscles (RBCs) lysis buffer treated cervical cell specimens. This second section is further divided in to two parts; the first part deals with normal and abnormal specimens, whereas later part explains the classification of treated normal, dysplastic and cancer specimens. The last section of this chapter discusses the influence of lymphocytes on the classification of cell specimens.

3.1 Total samples utilized in the study

Exfoliated cervical cell specimens were collected from patients visiting Tata Memorial Hospital, India, after obtaining informed and written consent. Hundred and ten exfoliated cell specimens were collected, out of which a total 5 and 11 cell specimens were excluded from the study due to the poor quality of spectra and inadequate samples, respectively.“Inadequate” specimen is terminology used by cytologist, referring to the specimens containing very less number of cell, such specimens are labeled as ‘inadequate’ – cell number are too low to provide any comment on specimen type. Therefore, the study was carried out using 94 exfoliated cervical cell specimens. Normal specimens were collected from subjects with non-cervical gynecological cancers having healthy cervix. Cancer specimens were collected from histopathologically certified cervical carcinoma patients. The study was approved by Institutional Review Board (IRB). Sample details are enlisted in Table 3.1.

Table 3.1: Total specimens utilized in the study.

Specimen details	No. of cases	Total no. of spectra	Abnormal	Normal
Inadequate specimens	11	0	7	4
Poor quality spectra	5	15	2	3
Utilized for analysis	94	498	49	45
Total	110	513	56	53

3.2 Exploring classification among untreated exfoliated cell specimens

3.2.1 Specimen details

The specimens were collected using cytobrush (HiMedia Laboratory Pvt. Ltd, India.) in 15 ml tube containing 1.5 ml of normal saline and transported at 4^0C. Exfoliated cells were then harvested by gently shaking the tubes containing cytobrush in normal saline. After this, the cytobrush was discarded and specimens were spun to obtain cell pellets. Thirty seven (37) cell specimens (17-normal and 20-cancers) were suspended in normal saline and centrifuged at 4000 rpm for 2 minutes to obtain pellets. Cell pellets were ready for Raman measurements.

3.2.2 Raman spectral acquisition details

A commercial Raman system HE-785 (Jobin-Vyon-Horiba, France) attached with superhead and objective (40x) was used in this study; this is photographically represented in Figure 3.1. The Raman spectra were acquired from cell pellets by placing them on the CaF2 window. A detail description of Raman system has been presented in chapter 2, section 2.3. Briefly, this system consist of a diode laser (Process Instruments) of 785 nm wavelength as excitation source and a high efficiency (HE-785) spectrograph coupled with a CCD (Synapse) as detection

element. The other component of the system, the 'Superhead', aids in optical filtering of unwanted noise including Rayleigh signals. Laser light delivery and Raman signal collection was carried out through the Superhead coupled with a 40X microscopic objective (Nikon, NA 0.65). The spectrograph of the Raman system has unmovable parts with fixed a 950 gr/mm grating. Spectral resolution as per manufacturer's specifications was ~4 cm^{-1} and the laser spot size at the sample was 4-5 µm. Spectra were integrated for 6-7 seconds and averaged over 3 accumulations. The laser power of 40 + 0.05 mW was kept constant during all the measurements.

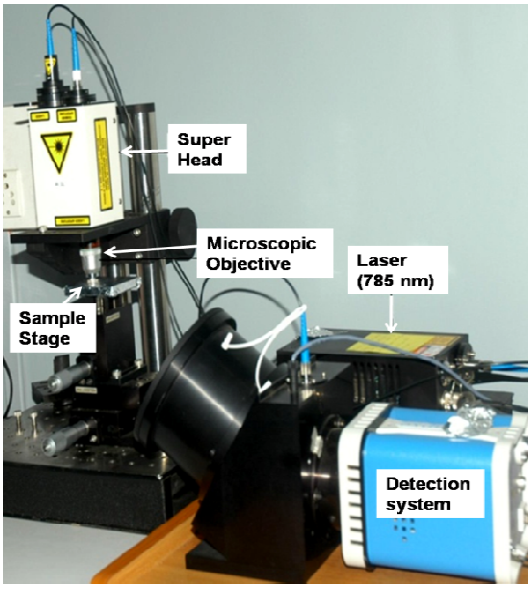


Figure 3.1: Pictographic representation of instrument used in the study.

3.2.3 Papanicolaou (Pap) staining

Sampling error may occur during exfoliation of cells from certified abnormal subjects, leading to false interpretations [11, 136]. To ensure cytological status, all the specimens were Pap stained after spectral acquisition. Pap staining of cells

was carried by the RAPID-PAP kit (Bio Lab Diagnosis Pvt. Ltd, India) as per the recommended protocol. The Pap smears were then subjected to two independent cytological examinations. Specimens were graded for red blood corpuscles (RBCs) and lymphocytes presence as mild, moderate and severe. All the specimens were categorized as normal and abnormal group. Abnormal specimens were also further divided into pre-cancer (HSIL, ASC-H, ASC-US) and cancer (SCC)

The protocol for Pap staining was as follows

1. Cells were smeared on a clean, ungreased slide and fixed in 100% of methanol solution.
2. Fixed smears were dipped in tap water for a minute and excess water was blotted out.
3. Smears were dipped for 45 seconds in RAPID-PAPTM nuclear stain.
4. Smears were washed in Scotte's tap water buffer for 30 seconds and excess water was blotted from the slide.
5. Smears were dipped for 30 seconds in RAPID-PAPTM dehydrant no.1 and then in no. 2 each.
6. Smears were dipped in for 45 seconds in RAPID-PAPTM cytoplasm stain.
7. Smears were washed in Scotte's tap water buffer and excess water was blotted from the slide.
8. Smears were dehydrated in a second bath of RAPID-PAPTM dehydrant for 30 seconds and air dried.
9. Smears were dipped in Xylene, dried and mounted with cover glass using a drop of Dibutyl Phathalate Xylene (D.P.X)

3.2.4 Raman spectral pre-processing

Preprocessing of Raman spectra was carried-out as per the previous discussed protocol in chapter 2, section 2.2.5. All the spectra were interpolated in the 900-1800 cm^{-1} range, first derivativized and this was followed by vector normalization and then subjected to PC-LDA.

3.2.5 Average spectra

Average spectra were computed as described in chapter 2, section 2.2.6. The baseline corrected, vector normalized spectra were used as representative of each group and to compute difference spectra.

3.2.6 Multivariate statistical analysis-PC-LDA

Data was analyzed by PC-LDA as described in chapter 2, section 2.2.7. The findings of PC-LDA were evaluated using LOOCV.

3.2.7 Raman spectral features

Thirty-seven cell specimens were subjected to Raman spectroscopy followed by Pap staining. A sum of 88 and 110 spectra were recorded from cell pellets of 17 and 20 certified normal and abnormal exfoliated cell specimens, respectively. Mean spectra of normal, cancer cell specimens and blood, along with their standard deviation are represented in Figure 3.2 A, B and C, respectively. Standard deviations exhibited minor intensity related variations within the groups. The average spectrum of exfoliated cell specimens from normal cases showed distinct bands at amide I (1660 cm^{-1}), δ CH_2 stretch (1450 cm^{-1}) and aromatic ring (1002 cm^{-1}) breathing of phenylalanine. The average spectra of abnormal cell

specimens exhibited strong features of C=C heme stretch (1620 cm^{-1}), fibrin (1570 cm^{-1}), δ CH_2 (1450 cm^{-1}), C-C symmetrical stretch (1234 cm^{-1}) from heme and Phenylalanine (1002 cm^{-1}), suggesting the presence of blood components like fibrin and RBC.

Differences in the amide I region were also observed in normal and abnormal spectra. Raman spectra of whole blood (Figure 3.2) were also acquired and corroborate spectral features of heme, fibrin in addition to a strong amide I (1660 cm^{-1}) was also observed. The Figure 3.2 (B and C) showed the prominent features of C=C heme stretch (1620 cm^{-1}), fibrin (1570 cm^{-1}), δ CH_2 (1450 cm^{-1}), C-C symmetrical stretch (1234 cm^{-1}) from heme and Phenylalanine (1002 cm^{-1}). This showed the similarities in abnormal and blood spectra present due common factor-blood. This suggested that differences among normal and abnormal specimens may arise due to blood presences.

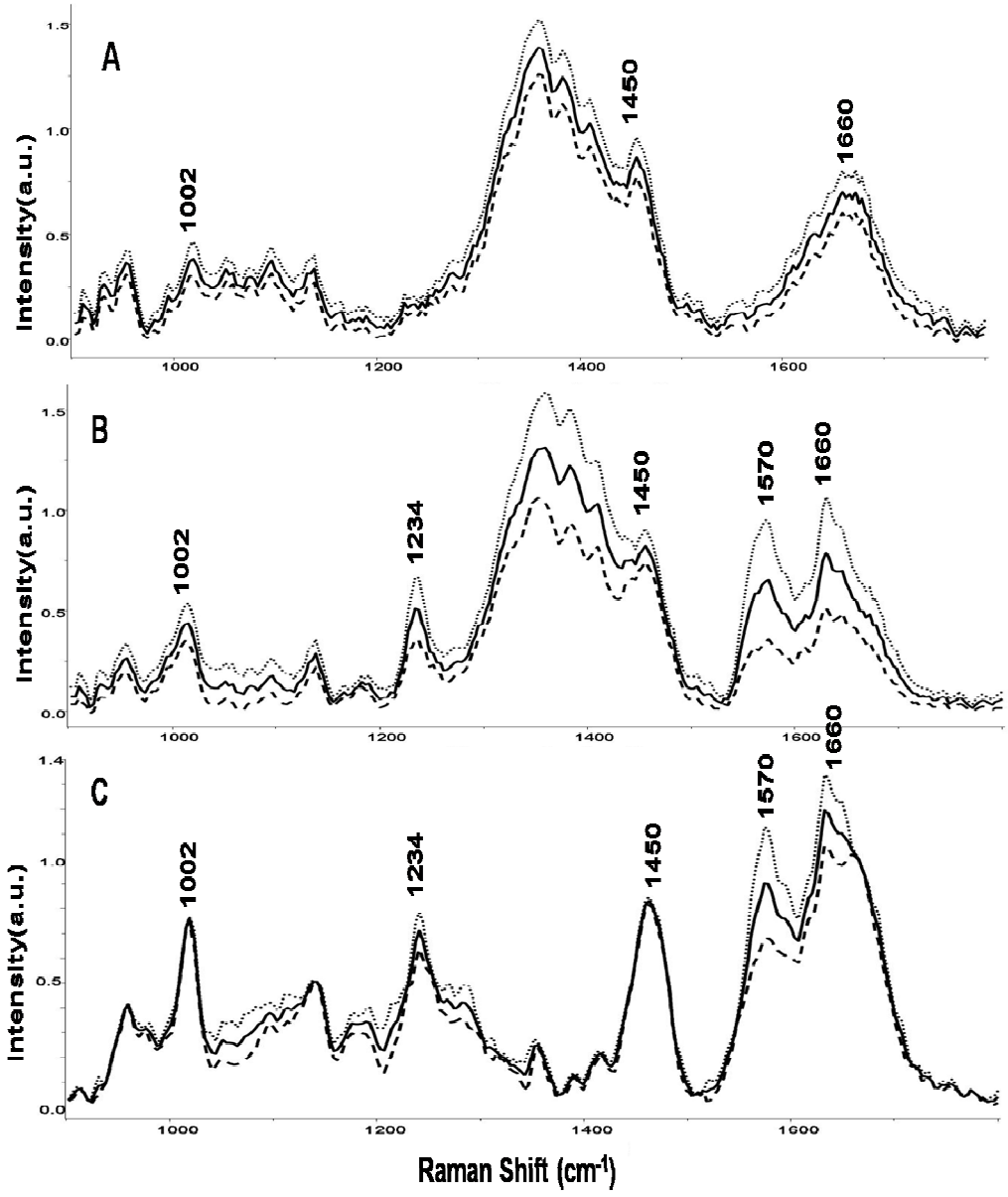


Figure 3.2: Mean spectra and standard deviation of (A) normal, (B) abnormal cell specimens and (C) blood.

Computation of a difference spectrum is one of the conventional methods to understand spectral differences over a selected spectral range and it can give information regarding moieties being altered. The mean normal spectrum was subtracted from the mean abnormal spectrum to compute difference spectra. The

positive peaks of difference spectra were from abnormal specimens while negative peaks belong to normal specimens (Figure 3.3). It was observed that positive peaks corresponding to C=C heme stretch (1620 cm^{-1}), fibrin (1570 cm^{1}), δCH_2 (1450 cm^{-1}), C-C symmetrical stretch of heme (1234 cm^{-1}) and phenylalanine (1002 cm^{-1}) were present. Features of the difference spectra were also suggestive of strong blood influence on spectra from abnormal specimens [71, 84, 137]. The negative peak (1660 cm^{-1}) belongs to normal specimens and suggests protein conformational changes were also observed.

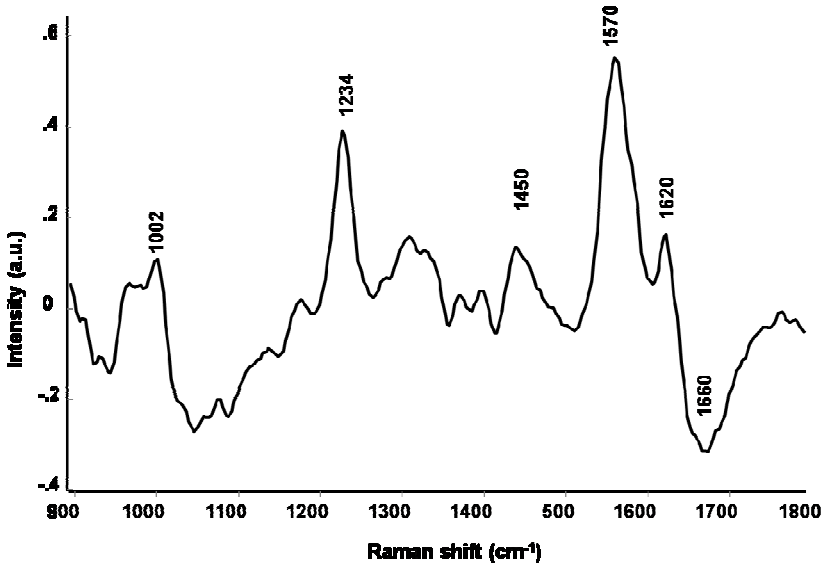


Figure 3.3: Difference spectra of abnormal-normal untreated exfoliated cell specimens

These observations were further established by staining cell samples used for Raman acquisitions. Presence of intact RBCs was also seen on stained slides of abnormal specimens which were absent in normal slides (Figure 3.6).

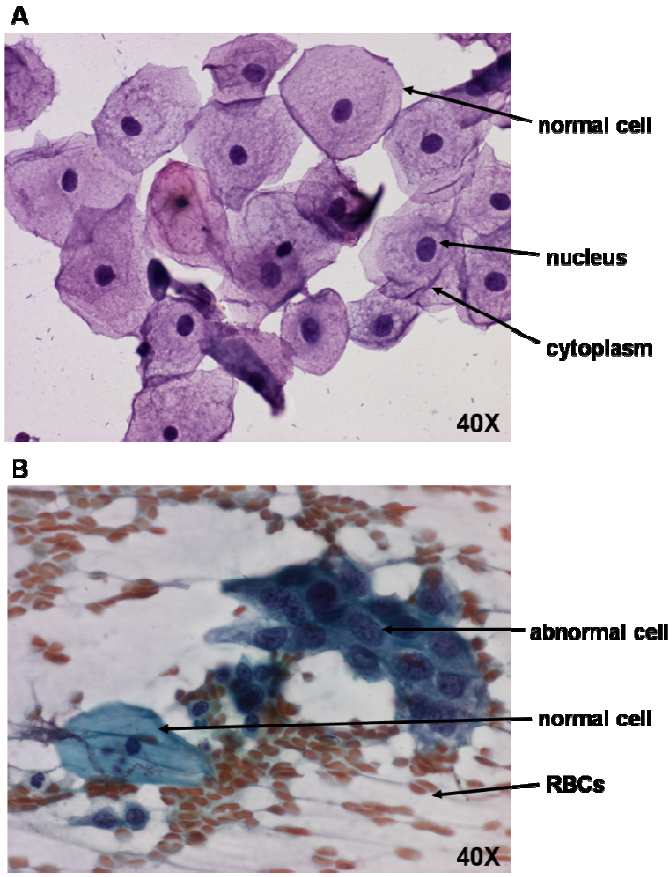


Figure 3.4: Pap stained visual image (40X) of: (A) normal smear (B) cancer smear, without treatment

3.2.8 Classification of the normal and abnormal untreated exfoliated cell specimens

PC-LDA was performed to explore the feasibility of classification of normal and abnormal specimens. As explained in chapter 2, it is method of classification that maximizes the ratio of inter-class variance to the intra-class variance in a data set thus resulting in maximal classification. The first five factors, giving ~ 90% of correct classification were selected for analysis (Figure 3.5A). The scatter plot shown in Figure 3.5 B, exhibits two clusters belonging to normal and abnormal group.

The PC-LDA results are also summarized in Table 3.2. It was observed that 80/88 normal spectra were correctly classified and 8 spectra were misclassified, whereas 99/110 spectra from cancers were correctly classified and 11 misclassified as normal. Classification efficiencies of 90 and 90.9 % for abnormal and normal specimens were observed, while in the validation step, 76/88 normal spectra and 93/110 abnormal spectra were correctly classified. Misclassification between normal and cancer can be attributed to presence of normal cells in abnormal specimens. LOOCV resulted in 84.5 and 86.4% classification efficiencies for abnormal and normal specimens, respectively.

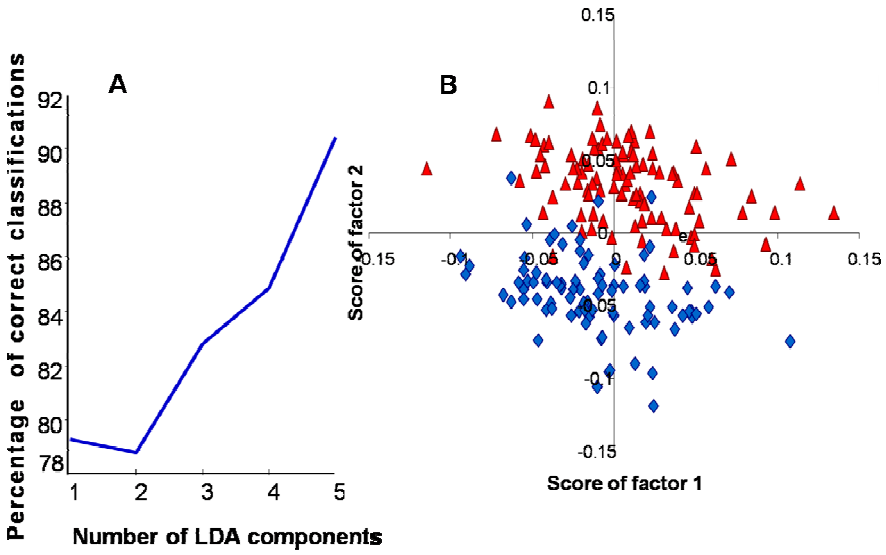


Figure 3.5: PC-LDA of normal and abnormal untreated exfoliated cell specimens (A) Scree plot (B) Scatter plot for PC-LDA normal smear (◆) cancer smear (▲).

In this study, the observed spectral features corresponding to heme and fibrin bands in cancerous specimens are indicative of blood as a contributing factor in classification. But, blood as a discriminating factor in cervical cancer may be misleading, since bleeding is a common occurrence during cervical

infections, uterine cancer and menstrual cycle. Therefore, it is pertinent to explore classification of pure cervical specimens i.e. devoid of blood influence. Hence, in the subsequent section a study in which specimens were treated with lysis buffer was carried out.

Table 3.2: PC-LDA of normal and abnormal untreated exfoliated cell specimens (A) Standard model (B) Leave one out cross validation.

A	**Standard model**			
	Abnormal	Normal	Total	Classification efficiency (%)
Abnormal	99	11	110	90
Normal	8	80	88	90.9
B	**Leave-one-out cross-validation**			
	Abnormal	Normal	Total	Classification efficiency (%)
Abnormal	93	17	110	84.5
Normal	12	76	88	86.4

3.3 Exploring the classification of RBCs lysis buffer treated exfoliated cell specimens

To circumvent the influence of RBCs on Raman spectroscopic classification of cervical exfoliated cell specimens, cell specimens collected from 57 subjects were treated with RBC lysis buffer. The Raman spectra were acquired from all the specimens and analyzed by PC-LDA. It is important to note that RBC lysis buffer solution includes ammonium chloride, which forms a mild osmotic pressure resulting in lysis of RBCs. As the membrane surrounding RBCs is weak, it leads to its rupture whereas, epithelial cells remain unaffected. By diluting the sample,

osmotic equilibrium is restored in order to avoid ill effects of prolonged exposure of epithelial cells [138]. As can be seen from Figure 3.6, no significant morphological changes in epithelial cells due to lysis buffer were observed. The structural morphology of treated and untreated cells (normal and abnormal) were correlative (Figure 3.4 and 3.6).

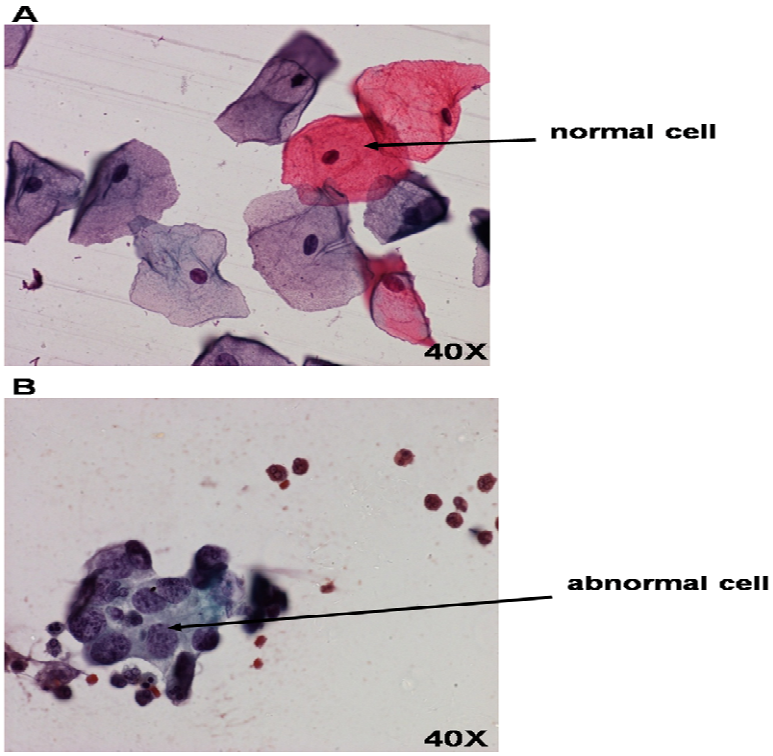


Figure 3.6: Pap stained visual image (40X), (A) normal smear and (B) cancer smear, post RBC lysis treatment

3.3.1 Specimen details

Fifty seven (57) cell specimens (28-normal and 29-cancers) were utilized in the study; specimens were centrifuged to obtain pellets as described earlier. The pellets obtained were then treated with 1 ml RBC lysis buffer for 15 min followed

by dilution with normal saline to stop the lysis reaction and centrifugation at 5000 rpm for 15 min to obtain pellets.

3.3.2 Raman spectral acquisition details

Raman spectra were acquired from the cell pellets by utilizing Raman HE-785 system (Jobin-Yvon-Horiba, France) as described in chapter 3, section 3.2.2. Spectral acquisition parameters were kept constant; briefly, Spectra were acquired for 6-7 seconds integration time and averaged over 3 accumulations. The laser power at specimen was 40 + 0.05 mW.

3.3.3 Pap staining

After spectral acquisition, cells were smeared on the glass slides before drying. Cell smears were immediately fixed in methanol and were subjected to Pap staining as earlier described in section 3.3.3. All the specimens were graded for the presence of RBCs, lymphocytes as mild, moderate and severe by two cytologists independently. The specimen's adequacy was also noted i.e minimum number of cell number required to comment on specimen's categorization. Specimens were further categorized as normal, HSIL, ASC-H, ASC-US or SCC.

3.3.4 Raman spectral pre-processing

Spectra were preprocessed as per the standard protocol, which is explained in detail in chapter 2, section 2.2.4.

3.3.5 Average spectra

Average spectra were computed as described in 3.2.5 section. Baseline corrected, normalized spectra were also employed to compute difference spectra.

3.3.6 Multivariate statistical analysis-PC-LDA

As mentioned in chapter 2, section 2.2.6. first derivative, pre-processed spectra were subjected to PC-LDA.

3.3.7 Raman spectral features

A vector normalized average of normal and abnormal cell specimens in the 900-1800 cm^{-1} region, post RBC lysis treatment is shown in Figure 3.7A and B, respectively. It was observed that the influence of blood on the spectra was removed to a great extent.

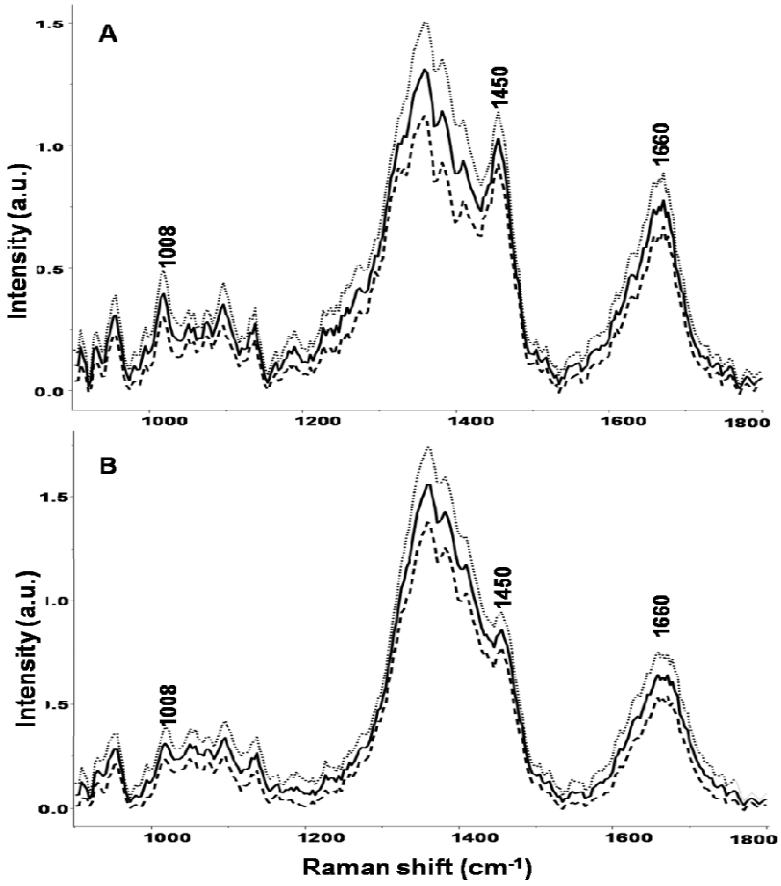


Figure 3.7: Mean spectra with their respective standard deviation of (A) normal specimen and (B) abnormal specimen post RBC lysis treatment.

This observation is corroborated by Pap stained slides (Figure 3.6). Variations in amide I (1660 cm^{-1}), amide III, δCH_2 (1450 cm^{-1}) and 1000–1200 cm^{-1} region were observed. The lack of heme and fibrin peaks suggests effective elimination of blood from the specimens. The improved Standard Deviation (SD) for treated specimens was observed as compare to untreated specimens. This observation may be due to the differences in RBCs contents in untreated specimens as compared to treated specimens, which were devoid of RBCs (Figure 3.2 and 3.6).

Difference spectrum was computed by subtracting average normal spectrum from average abnormal spectrum (Figure 3.8). Bands at 1660, 1450 and 1006 cm^{-1} indicated increase in protein as well as changes in secondary structure of proteins designated by a positive amide III peak were observed.

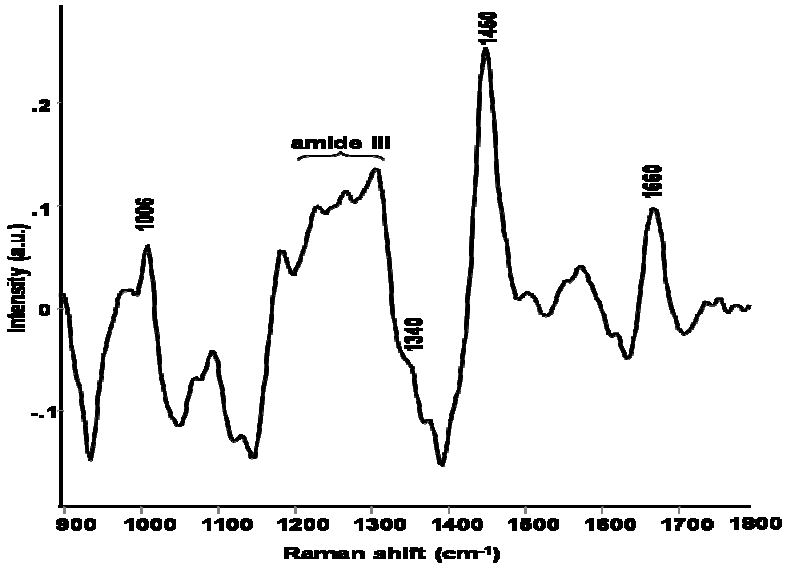


Figure 3.8: Difference spectra abnormal-normal untreated specimens.

3.3.8 Classification of normal and abnormal smears

To explore the classification of post RBC lysis treatment exfoliated cervical cell specimens, 5 factors, contributing ~79% of correct classification, were selected for PC-LDA (Figure 3.9A). The scatter plot exhibited two separate clusters with overlap corresponding to normal and abnormal cell, respectively, which is shown in Figure 3.9B.

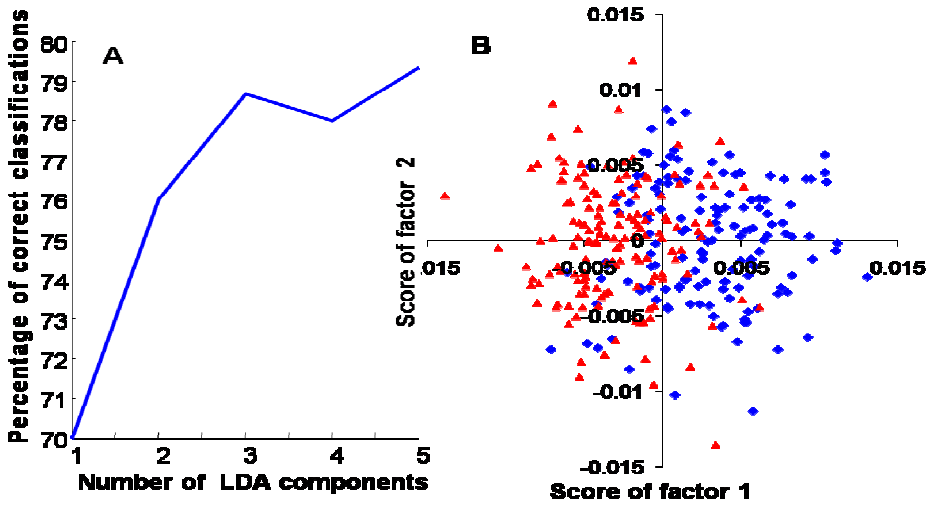


Figure 3.9: PC-LDA of normal and abnormal RBC lysis treated exfoliated cell specimens (A) Scree plot (B) Scatter plot for PC-LDA: normal (🔷) abnormal (🔺).

The PC-LDA results are summarized in Table 3.3. For the standard model, it was observed that 119/150 (79.3%) abnormal and normal spectra were correctly classified, while the remaining spectra were misclassified into the other groups (Table 3.3A). The confusion matrix for LOOCV demonstrated 118 /150 (78.7%) normal as well as 119/150 (79.33%) abnormal spectra were correctly classified whereas, 32/150 abnormal and 31/150 normal spectra were misclassified (Table 3.3 B).

Table 3.3: PC-LDA of normal and abnormal RBCs lysis treated exfoliated cell specimens (A) Standard model (B) Leave one out cross validation.

A	**Standard model**			
	Abnormal	Normal	Total	Classification efficiency (%)
Abnormal	119	31	150	79.3
Normal	31	119	150	79.3
B	**Leave-one-out cross-validation**			
	Abnormal	Normal	Total	Classification efficiency (%)
Abnormal	118	32	150	78.7
Normal	31	119	150	79.33

Sample heterogeneity, in the context of varying numbers of normal and abnormal cells in 'abnormal' specimens may be the reason of the observed misclassification between normal and abnormal group. For instance, the distribution of abnormal cells in samples utilized in the study ranges between 1 to 40 %. Also, it is very significant to note that Raman spectra were measured with the laser spot of 5-10 μm and the depth of penetration could be around 40 μm from a thick cell pellet. Furthermore, as per manufacturer specifications, the probing volume was ~ 500 cubic microns. Since a pile of cells represent pellet, the probe section or sample could be at various dissimilar cellular components and many cells resulting in co-localization of cancer cells. It is important to note that, after removal of blood, although abnormal samples were contaminated by normal cells, the classification efficiency was ~80%, which is analogous to the Pap test. This suggests the possibility of classification of normal and abnormal Pap specimens using Raman spectroscopy.

High classification efficiency (Table 3.3) was observed for untreated samples but it might lead to misleading interpretations or results. In such situations, regardless of significantly lower specificity/sensitivity, RBC treatment is a superior approach. Even though the results were equivalent to conventional Pap test, the efficacy of this approach can be further improved by developing more robust models. More robust models can be built by selectively accruing abnormal specimens with an extremely high number of abnormal cells, therefore reducing the dominance of normal cells in abnormal specimens consecutively and reducing bias in classification. If such models are developed, test spectra can be evaluated against the model and specimens in which, single spectrum matches with cancer, it can be assigned as cancer. This is conventional standard practice in histopathology or cytology examination where several sections are examined and even if one slide show a patch of malignant cells the specimen is treated as cancer. Therefore, potential Raman spectroscopic studies on pure cancerous and precancerous specimens to build true standard models and validation by huge blinded specimens can further establish the role of RS as an important cervical cancer screening tool into clinics.

3.3.9 Classification of normal, dysplastic and cancer exfoliated cell specimens

To investigate the feasibility of differentiation of normal, pre-cancer and cancer exfoliated cell specimens, data was analyzed using Principal component analysis (PCA). Preprocessed spectra were subjected to data analysis by using PCA algorithms implemented in in-house build software. A scatter plot for PCA of normal, precancerous and cancer exfoliated cell specimens is shown in Figure

3.10. Three clusters, belonging to normal, precancer and cancer exfoliated cell specimens were observed. The findings indicate the tendency of classifications. However, the overlap between these clusters of normal, precancer and cancer cells was also observed. This may be due to specimen heterogeneity.

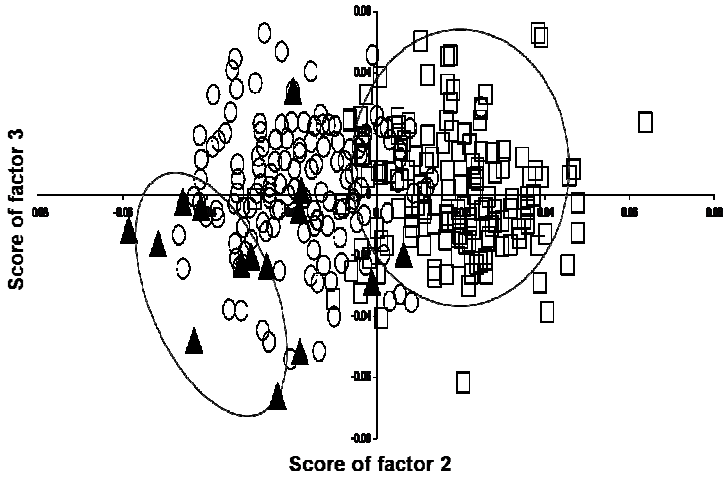


Figure 3.10: Scatter plot for PCA of normal (□), dysplastic (○) and cancer (▲) specimens.

3.4 Exploring the influence of lymphocytes on classification of normal and abnormal exfoliated cell specimens

RBCs influence the classification of exfoliated cervical cell specimens; as lymphocytes are the one of constituent of blood, they may be another confounding factor on the classification of smears, post RBC removal. Hence the influence of lymphocytes on the classification of exfoliated cell specimens was explored

3.4.1 Lymphocyte extraction

Lymphocytes were isolated from heparinized blood of healthy adult volunteers by density gradient sedimentation using ficoll/isopaque [139]. Peripheral blood collected in heparin (Sigma, USA; 100 U/ml) was diluted with equal volume of normal saline (0.82% NaCl in double distilled water) 10 ml of diluted blood was loaded as 2.5 ml of ficoll-Hypaque [24 parts of 9% ficoll 400 (Sigma, USA) + 10 parts 33.3% sodium diatrizoate (Sigma, USA), specific gravity to 1.077 + 0.001] and centrifuged at 1,500 rpm for 20 min at room temperature (RT) using a swing-out rotor. Lymphocytes were collected from the interface between ficoll hypaque and plasma. Cells were washed thrice with sterile normal saline.

3.4.2 Specimen details

Exfoliated cervical cell specimens were collected from 10 subjects having healthy cervix in normal saline, with no clinical history of abnormal Pap test. Specimens were spun to obtain cell pellets which were further pulled to single pellets.

3.4.3 Exploring the influence of lymphocytes on classification of normal and abnormal exfoliated cell specimens

To evaluate the influence of lymphocytes on classification of exfoliated cell specimens, the experiment was designed, in which lymphocytes were mixed in different ratios with exfoliated cell specimens. Cells were suspended in saline solution and counted using a Neubauer chamber. Lymphocytes were then mixed with exfoliated cervical cells in various ratios such as 1:0, 1:1, 1:2, 1:3 and 0:1 (Figure 3.11).

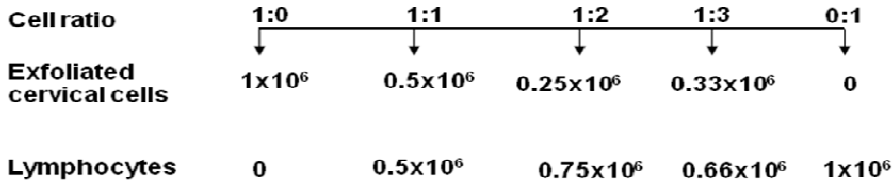


Figure 3.11: Schematic representation of experiment: mixing different ratio of lymphocytes to exfoliated cervical cells.

The total cell number was kept constant. Cells were spun to obtain cell pellets.

a. Raman spectral acquisition details

Raman spectra of cell pellets were recorded by Raman HE-785 system as described in section 3.2.2. Spectra were measured for 7 seconds acquisition time and averaged over 3 times. Laser power at specimen was 40+ 0.05 mW during all the measurements.

b. Pap staining

Subsequent to spectral measurements, cells were spread on glass slides, fixed on slides and were subjected to Pap staining as described in the section 3.3.3. All the specimens were graded for the presence of lymphocytes as mild, moderate and severe.

c. Raman spectral pre-processing and data analysis

Spectra were preprocessed as described in section 2.3.4 [128-130]. In summary, spectra were corrected for CCD response, background signals, first derivatization, interpolation and vector normalization. All the pre-processed Raman spectra were subjected to Principal Component Analysis (PCA).

The scatter plot for PCA is exhibited in Figure 3.12. It was observed that two exclusive clusters belonging to lymphocytes and exfoliated cell + lymphocyte were observed. Overlap was observed between the exfoliated cell specimens, exfoliated cell specimen + lymphocyte spectra. The cytological mixture of cells with 1:1 ratio were grade by cytologist as mild where as 1:2, 1:3 were graded as moderate/ severe. It was observed that Raman spectroscopy can detect the moderate and severe kind of inflammatory cells present in the specimens.

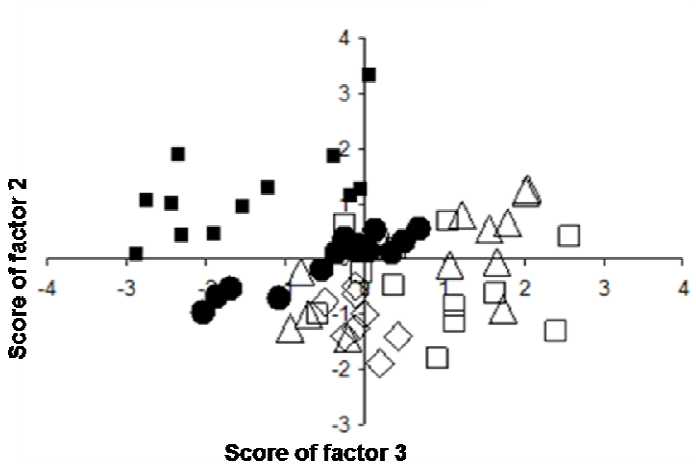


Figure 3.12: Scatter plot for Principal component analysis of normal exfoliated cell specimens, mixture of exfoliated cells and lymphocytes in 1:0 (●), 1:1(◊), 1:2 (□), 1:3(Δ) and 0:1 (■), respectively.

3.4.4 Exploring the lymphocytes influence on classification of exfoliated cell specimens utilized in the study.

It has been reported that lymphocytes are aggregated at the site of a tumor [140]. Hence, they can influence the classification of exfoliated specimens. To evaluate the influence of lymphocytes on the classification of exfoliated cell specimens, the cytological categorization of all the specimens used for Raman spectroscopy as mild (+), moderate (++) and severe (+++) was carried out. As blood played a

role in the classification of exfoliated cell specimens, RBCs lysis buffer treated cell specimens were used. Table 3.4 shows the pattern of lymphocytes presence in normal and abnormal cell specimens. Among abnormal specimens, 3/29, 11/29, 14/29 and 1/29 exhibited mild, moderate, severe and absence of lymphocyte, respectively. Whereas, normal specimens showed 2/28 lymphocyte absence, 13/28, 11/28 and 2/28 exhibited mild, moderate and severe, respectively. It was observed that, in the specimens used in our study, there was almost equal distribution of lymphocytes in normal and abnormal exfoliated cell specimens. Hence, the observed classification in treated exfoliated cell specimens was not due to lymphocyte.

Table 3.4: Distribution of lymphocytes in normal and abnormal exfoliated cell specimens- negative = absence of lymphocytes, + = mild, ++ = moderate and +++ =severe lymphocyte presence.

	-ve	**+**	**++**	**+++**	**total**
Abnormal	1	11	14	3	29
Normal	2	13	11	2	28

3.5 Summary

The utility of *ex vivo* Raman spectroscopic approach to differentiate exfoliated cervical cell specimens and the influence of diverse factors on its classification were explored. This chapter is divided into three sections: the first section describes about Raman spectroscopic classification of untreated specimens into normal and abnormal categories. The second section explains about RBCs influence on Raman spectroscopic classification of cervical cell specimen. This

section is further divided in to two parts, first part discuss about the classification of normal and abnormal RBCs lysis buffer treated specimens whereas later part elucidate about further classification of this treated specimens in to normal, dysplastic and cancer categories. The third section of this chapter discuss about the influence of lymphocytes on the classification of exfoliated cell specimens. To date, to the best of our knowledge, no studies toward classification of normal and abnormal specimens using cytological certified cells specimens by Raman Spectroscopy have been reported.

The work presented in this chapter has been summarized as follows:

- ✓ The PC-LDA yielded classification efficiencies of 86% and 84% for normal and abnormal specimens in untreated exfoliated cervix cell specimens, respectively.

- ✓ Most of the cervical cancer subjects bleed due to the high vascular nature of tumors. The presence of RBCs in specimens can also occur in non cervical cancerous conditions. Hence, the influence of RBCs on the classification of exfoliated cervix cell specimens was explored. PC-LDA resulted in classification efficiencies of 79% and 78% for normal and abnormal smears, respectively. Misclassifications in both the approaches can be attributed to the predominance of normal cells in abnormal specimens. Although higher classification efficiency was observed for untreated samples, it might lead to false interpretations and misleading results because bleeding can occur in other normal conditions. In this

context, despite lower specificity/sensitivity, RBC lysis treatment may be a better approach.

- ✓ The classification of normal, precancerous and cancerous exfoliated cell specimens was also explored. It was observed that cancerous and normal spectra showing the tendency toward classification, whereas precancerous spectra showed overlap with normal and cancerous.

- ✓ The influence of lymphocytes on the classification of exfoliated cervical cell specimens was also explored. The PCA findings suggest that the presence of lymphocytes in lower concentrations had minimal or no influence on the classification of cell specimens.

4 HPV EXPRESSING AND NON EXPRESSING CERVICAL CELL LINES

Introduction

Human papillomavirus (HPV) is one of major etiological factors for cervical cancer, the second most common malignancy among women worldwide [5, 132]. The HPV is a group of more than 150 type of DNA viruses belong to the papillomavirus family. Among all subtype, HPV-16 and HPV-18 are high risk type, associated with approximately 70% of cervical cancers [5]. HPV-16 is commonly associated with squamous cell carcinomas, HPV-18 with adenocarcinomas. High-risk strains mainly infect the epithelium to promote proliferation, leading to uncontrolled proliferation of cells. The high-risk HPV strain contains E5, E6 and E7 oncogenes resulting in cell damage and abnormal cell proliferation by cooperatively interfering with the functions of p53 and pRb, the cellular tumor suppressor proteins [5]. It is also known that persistent HPV infection is required for the development of cervical cancers [5].

Infection with high-risk HPV is the key risk factor for cervical cancer. Thus, HPV testing has been included in the range of clinical options for cervical cancer screening [132]. Detection of viral DNA is the basis for HPV testing. HPV testing is known to have some limitations, as it is expensive, time-consuming and requires sophisticated laboratory infrastructure. The aim of the study is to evaluate the ability of Raman spectroscopy to detect HPV induced cellular differences in the cervical cancer cell lines. HPV 18 positive HeLa, HPV-16 positive SiHa and HPV negative C33A cell lines were used for the study. The findings of the study are discussed in this section.

4.1 Materials and methods

4.1.1 Cell lines

HeLa: HPV-18 positive, human cervical adenocarcinoma cell line (CCL-2, ATCC, USA)

SiHa: HPV-16 positive, human cervical squamous- carcinoma cell line (HTB-35, ATCC, USA)

C33A: HPV negative, human cervical carcinoma cell line (HTB-31, ATCC, USA)

4.1.2 Culture medium

IMDM medium (Invitrogen Life-Technologies, Grand Island, N.Y) medium powder was dissolved in deionised water and supplemented with sodium bicarbonate (SRL, Ranbaxy Ltd, India) and, HEPES buffer (Sigma, St Louis, MO) as per manufacturer's instruction. The medium was sterilized by membrane filtration (0.45µm, Millipore Co, Bedford, MA).

Complete medium was prepared by supplementing IMDM with 10% inactivated fetal calf serum, FCS; Invitrogen Life-Technologies, Grand Island, N.Y), penicillin (100 IU/ml; Alembic Chemicals, India), streptomycin (100 µg/ml; Alembic Chemicals, India), mycostatin (5µg/ml; Sigma, USA), gentamycin (40 µg/ml; Schering corpa, India) L-glutamine (2 mM; HiMedia, India) and β-mercaptoethanol (5 x 10-5 M, Sigma, USA).

4.1.3 Cell culture

HeLa, SiHa and C33A cell lines were grown in IMDM (Invitrogen). Cell lines were incubated at 37˚C in 5% $C0_2$ and cultured to 70-80% confluence.

4.1.4 Sample preparation for Raman spectroscopy

Cells were detached from the flask using 3% trypsin-EDTA (Sigma-Aldrich) solution, incubated until cells detached from the surface, followed by addition of IMDM containing foetal bovine serum (FBS). Cell suspensions were centrifuged at 1200 rpm for 10 minutes to obtain pellets. Cell pellets were washed twice with phosphate buffer saline (PBS) and centrifuged at 1200 rpm for 10 minutes after each wash. Finally, the supernatant was removed and cell pellets were transferred onto CaF_2 window. The number of cells per pellet was adjusted to 1 million cells so as to keep uniformity in experiment. The dimensions of the cell pellets were ~ 4 x 4 x 2 mm.

4.1.5 Raman spectral acquisition details

Cell pellets were placed on CaF2 window and spectra were recorded using a HE-785 commercial Raman spectrometer (Jobin-Yvon-Horiba, France). This system is described in detail in chapter 3, section 3.3.2. Spectra were integrated for 6 seconds and averaged over 3 accumulations. The laser power at the specimen was 40+ 0.05 mW. Approximately 8-9 spectra per pellet were acquired. Reproducibility was confirmed by three independent experiments.

4.1.6 Average spectra

Average spectra were computed as described in chapter 2, section 2.2.4.

4.1.7 Raman spectral pre-processing

Raman spectra were preprocessed as described in chapter 2, section 2.2.4. Preprocessing steps for cell lines are shown in Figure 4.1. Analysis of the first derivative, pre-processed spectra was carried out using Principal Component Analysis (PCA).

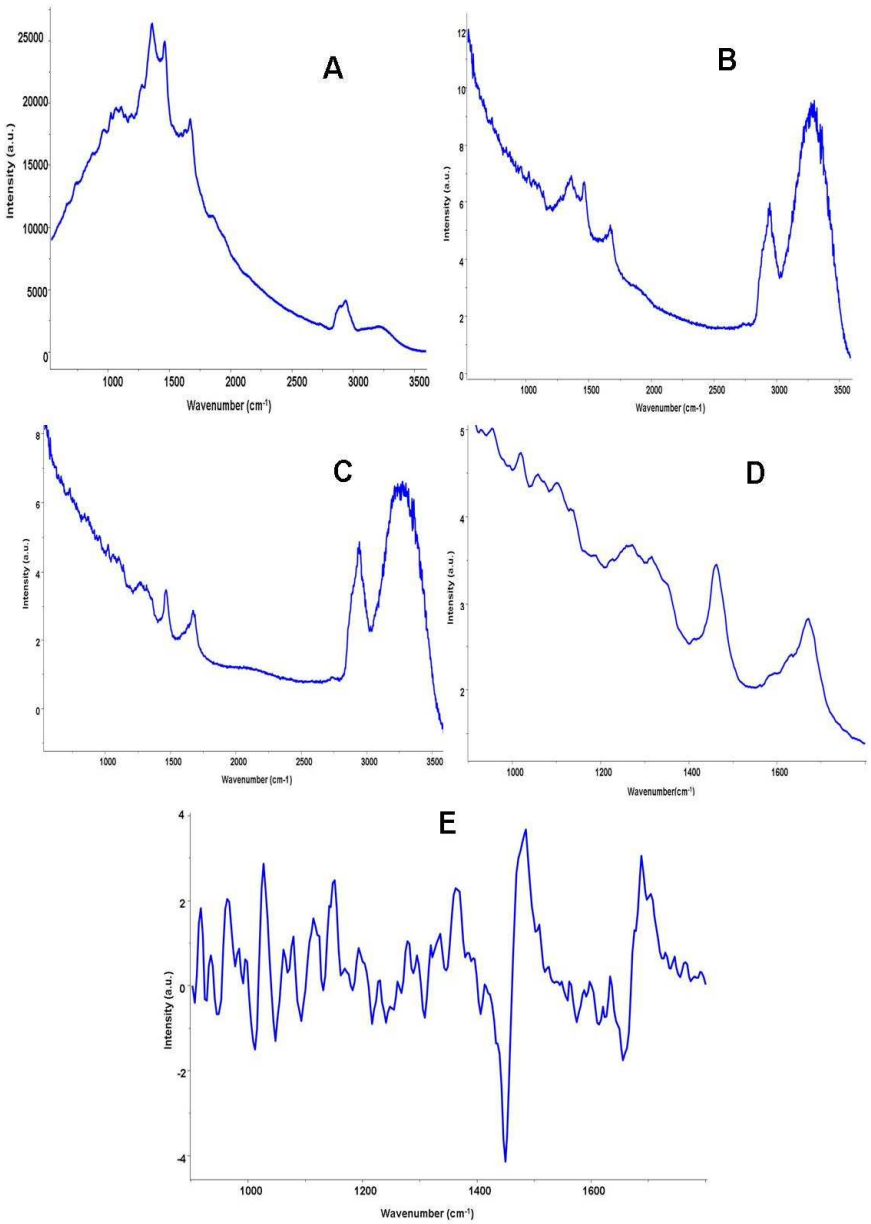


Figure 4.1: Representative cell line spectra at different pre-processing steps. A. Raw spectrum B. CCD response corrected C. Background corrected D. interpolated and E. First derivative.

4.2 Results and discussion

4.2.1 Raman spectral features

The goal of the present study was to identify the HPV induced cellular differences in cervical cancer cell lines utilizing Raman spectroscopy. The Raman spectra of HPV-18 positive HeLa, HPV-16 positive SiHa and HPV negative C33A cell lines were acquired. Vector normalized average Raman spectra of C33A, HeLa and SiHa cells, along with their standard deviations are shown in Figure 4.2 A, B and C, respectively. To understand spectral heterogeneity within the group, standard deviations were also computed. Minor intensity related variations among the group were observed. As is evident from Figure 4.2, contributions of protein, lipid, DNA and amino acids were observed in the mean spectra of all groups. The annotations are in good agreement with the observations made by Ostrowsksa *et al.* [118]. HPV positive cells, HeLa and SiHa showed highly intense bands at amide I (1660 cm^{-1}), δCH_2 (1550 cm^{-1}) and band at 1340 cm^{-1}. The high protein and nucleic acid signals in HeLa and SiHa cells may be due to HPV induced changes cause because of increased cell proliferative rates. Variations in vibrations connected with nucleic acid (1340 and 1098 cm^{-1}) were also observed. Further differences in amide I and amide III bands between HPV expressing and non-expressing cells were also observed. Differences associated with protein and nucleic acid composition support the earlier reports [117, 118], which demonstrated that HPV genome integration in host chromosome causes subsequent increase in the cell proliferation rate, causing cells to turn malignant.

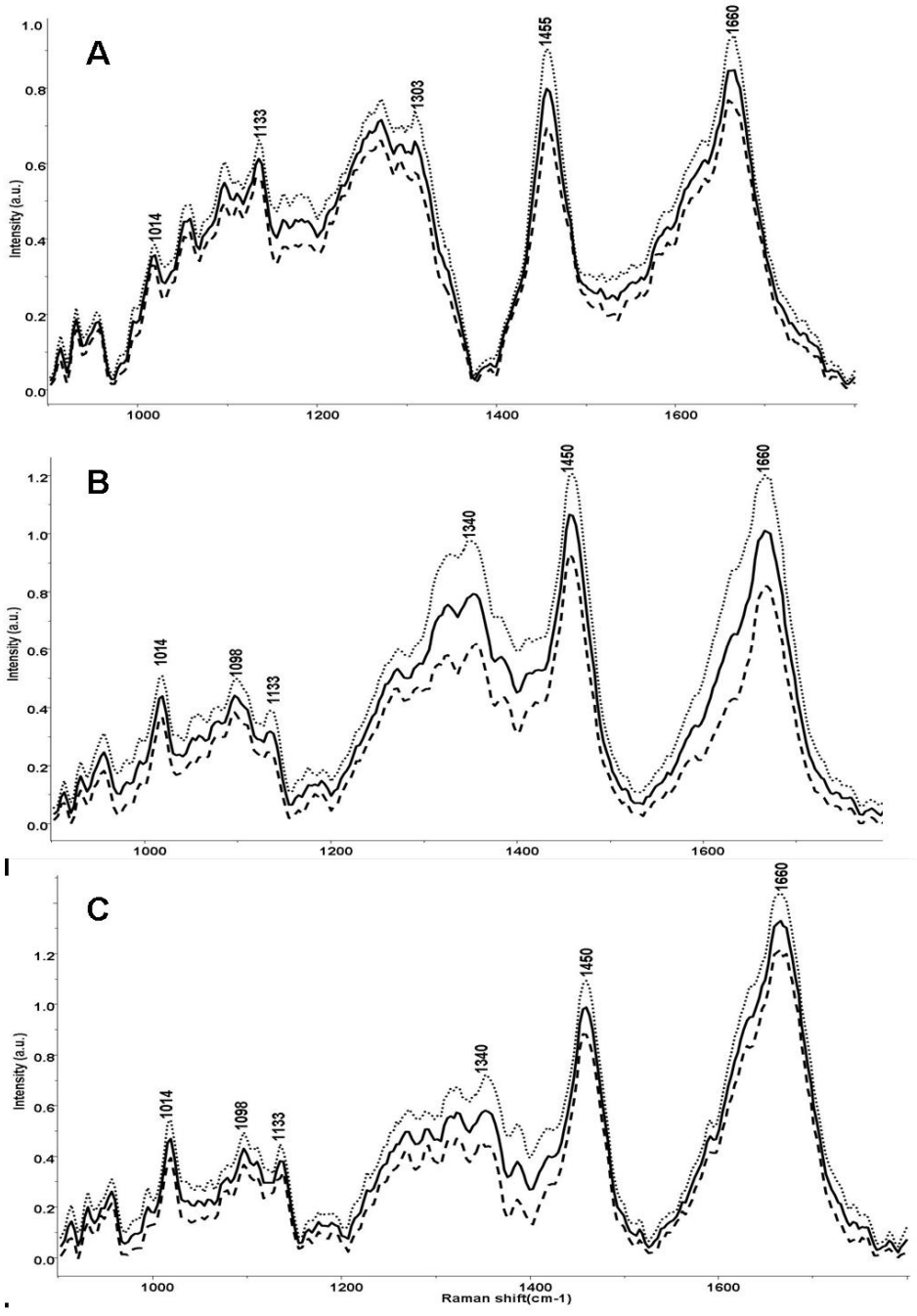


Figure 4.2: Mean spectra with their standard deviation (A) C33A (B) HeLa (C) SiHa cells.

To explore further the spectral variations between HPV expressing and non-expressing cells; difference spectra were computed by subtracting the average spectrum of C33A from HeLa and SiHa spectra. Difference spectra (HeLa-C33A, SiHa-C33A and HeLa-SiHa.) are shown in Figure 4.3 A, B and C, respectively.

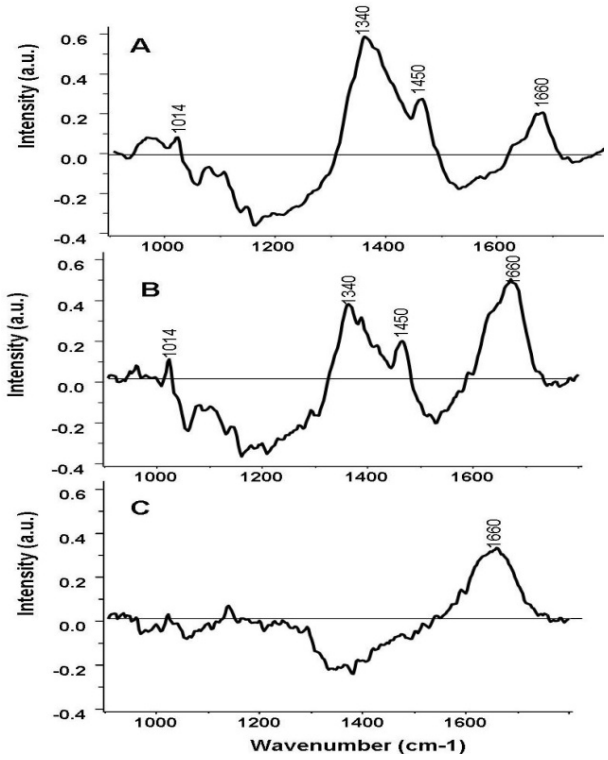


Figure 4.3: Difference spectra (A) HeLa- C33A (B) SiHa- C33A (C) HeLa-SiHa.

The positive peaks of the difference spectrum are from the average HPV expressing cell spectrum (HeLa or SiHa) and negative bands are due to the C33A cells spectrum. Positive peaks of proteins like amide I (1660 cm^{-1}), δCH_2 deformation (1450 cm^{-1}), and DNA (1340 cm^{-1}) were observed in HPV expressing cells, which were comparatively weak in C33A cells. These observations were consistent with the observations made by earlier studies [117, 118]. Ostrowsksa *et al*. reported that HPV negative C33A has no HPV copy per cell, HPV-18 positive

has 20-50 integrated HPV copies per cells, HPV-16 positive SiHa contains 1-2 integrated HPV strands.

The Figure 4.3 C, represents the difference spectrum of HeLa- SiHa exhibiting minor protein related changes in the spectra. As can be seen in Figure 4.3C, presence of amide I (1660 cm^{-1}) indicating high amount of proteins in HeLa cells as compared to SiHa corroborating earlier study by Ostrowsksa *et al*. However, the observed differences due to the different cell type cannot be ruled out [141].

4.2.2 Multivariate statistical analysis

To investigate the feasibility of differentiation among HPV expressing and non expressing cell lines, Principal components analysis (PCA) was used. For visual discrimination, each of the spectra in the newly formed co-ordinate space of selected PCs was projected. Preprocessed, first derivative spectra were subjected to data analysis by using PCA. Profiles of PCs or factor loadings can provide vital clues on biochemical dissimilarities among different classes. The first three and two significant discriminating PCs were selected for 3D and 2D visualization of data, respectively (Figure 4.5 and 4.6). The spectral variability observed in the difference spectra is corroborated by the loading plots, suggesting variations in protein and nucleic acid content of HPV expressing and non-expressing cells.

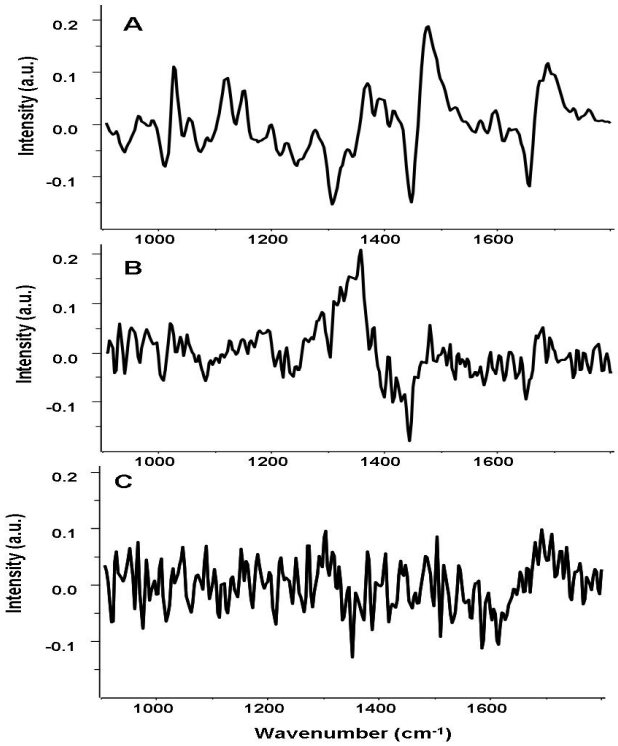
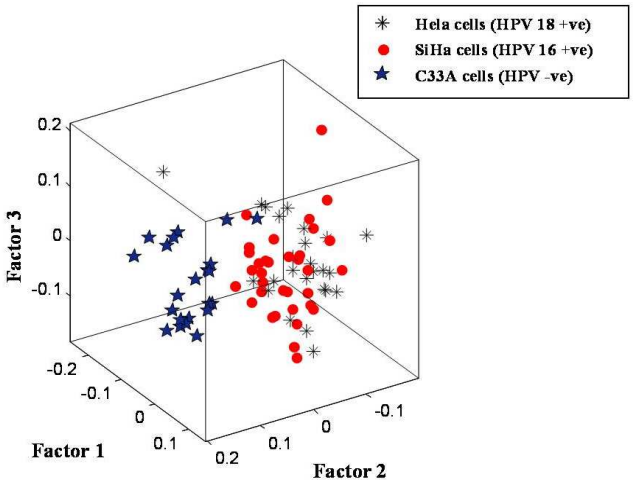


Figure 4.4: PCA analysis for HeLa, SiHa and C33A cell line (A) Loading of factors 1(B) Loading of factor 2 and (C) Loading of factor 3.

The first PC has four major bands that correspond to amide I, δCH_2, 1340 and 1014 cm^{-1} and the second PC has two main bands, at 1660 and 1340 cm^{-1}, corresponding to protein and nucleic acid contributions to classification. The cumulative variance of 60%, 75% and 82% was provided by PCs 1, 2 and 3 respectively (Figure 4.4).

The 3D and 2D scatter plot for PCA is shown in Figure 4.5 and 4.6 respectively. Two clusters belonging to HPV positive and HPV negative cells were observed. Overlap between the clusters of HeLa and SiHa cells was observed. This indicates that cells of these two populations, HPV expressing cells, have similar molecular profile, including subtle variations. However, these

populations of cells (HPV positive) were clustered apart from the HPV non-expressing cells (C33A), exhibiting their differences from HPV negative cells. The findings corroborate earlier reports that Raman spectroscopy can distinguish HPV expressing and non-expressing cells [117, 118].



4.5: 3D scatter plot for Principal Component Analysis for HeLa(*), SiHa (●) and C33A (★) cell line.

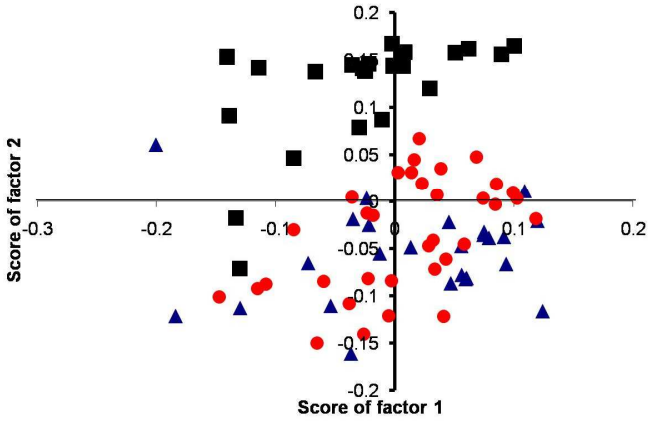


Figure 4.6: 2D scatter plot for Principal Component Analysis: HeLa(▲), SiHa(●) and C33A (■) cell line

4.3 Summary

The chapter aims to evaluate the ability of Raman spectroscopy to detect HPV induced cellular differences in the cervical cancer cell lines. HPV 18 positive HeLa, HPV-16 positive SiHa and HPV negative C33A cell lines were used for the study. Mean and difference spectra exhibited variations associated with protein and nucleic acid composition. The PCA scatter plot showed two clusters belonging to HPV positive and HPV negative cells. Overlap between the clusters of HeLa and SiHa cells was observed, indicating similarity among them. The findings suggest that Raman spectroscopy can distinguish HPV expressing and non-expressing cells.

5 SUMMARY

The work presented in the thesis describes the utility of Raman spectroscopy in conjunction with multivariate statistical tools for the improved diagnosis of cervical cancer, both *in vivo* and *ex vivo*. Specifically, it aims to evaluate the efficacy of Raman spectroscopic methods for non-invasive/minimal-invasive and objective screening/diagnosis of cervical cancers. The major highlights of the work are as follows:

5.1 *In vivo* Raman spectroscopy of cervical cancers

A fiberoptic Raman system for *in vivo* cervical cancer applications was procured and standardized. To standardize the data acquisition, analysis as well as the spectral reproducibility, spectra of *ex vivo* normal and tumor cervix tissues were acquired. It was observed that the normal cervical tissue spectra were rich in collagenous type of proteins, while tumor tissue spectra were dominated by non-collagenous type proteins and nucleic acid. It was observed that PC-LDA standard models of normal and tumor tissue spectra exhibited prediction efficiencies of 94 and 91 %, respectively. The misclassification between both groups can be primarily attributed to the tissue heterogeneity.

The feasibility of acquiring good quality *in vivo* Raman spectra under clinically implementable time in Indian population was demonstrated. A total of 442 spectra were acquired from 103 subjects. Mean normal cervix spectra showed collagenous type of proteins while tumor tissue spectra were dominated by non-collagenous type of proteins and nucleic acid. The feasibility of objectively classifying tumor and normal cervix *in vivo* sites was tested. Prediction efficiencies of 94.5 and 100 % for normal and tumor *in vivo* sites were observed,

respectively. The finding suggests the applicability of Raman spectroscopic methods for objective, noninvasive and rapid cervical cancers diagnosis and corroborates earlier studies.

To circumvent the influential parameters like menopausal status, hormonal status, age, and parity on the classification of the data, the utility of normal vaginal sites as an internal control was also explored. Multivariate statistical analysis of normal cervix and vaginal sites of tumor and normal subjects was carried out. The findings suggest that vaginal sites can be used as internal control, where the normal cervix sites may be unavailable due to advancement of disease. This will also help to circumvent the inter-patient variability caused due to differences in age, parity, hormonal and menopausal status.

5.2 Raman spectroscopic study on exfoliated cervical cell specimens

To the best of our knowledge, no studies toward classification of normal and abnormal specimens using certified exfoliated cervical cells specimens with Raman spectroscopy have been reported.

Raman spectroscopic studies on the classification of normal and abnormal exfoliated specimens were carried out and the effect of factors like presence of RBCs and lymphocytes on their classification was studied. The classification of untreated exfoliated cervix cell specimens was explored. PC-LDA yielded classification efficiencies of 86% and 84% for normal and abnormal specimens, respectively.

Most cervical cancer subjects bleed due to the high vascular nature of tumors. The presence of RBCs in a specimen can also occur in non cervical cancerous conditions. Hence, the influence of RBCs on the classification of exfoliated cervix cell specimens was explored. PC-LDA resulted in classification efficiencies of 79% and 78% for normal and abnormal RBCs lysis treated cell specimens, respectively. Misclassifications in both the approaches can be attributed to the predominance of normal cells in abnormal specimens. Even though higher classification efficiency was observed for untreated samples, it might lead to misleading results. Since, bleeding is a common occurrence during cervical infections, uterine cancer and menstrual cycle. Hence, RBC lysis treatment of exfoliated cervical cells may be a better approach.

The classification of normal, precancerous and cancerous exfoliated cell specimens was also explored. It was observed that precancerous spectra overlapped with normal and cancerous, whereas cancerous and normal spectra showed the tendency toward classification.

The influence of lymphocytes on the classification of exfoliated cervical cell specimens was studied. The PCA findings suggest that the presence of lymphocytes in lower concentrations had minimal or no influence on the classification of exfoliated cell specimens.

5.3 Raman spectroscopic study of HPV positive and negative cell lines

The High Risk Human papillomavirus (HR-HPV) is one of major etiological factors for cervix cancer; HPV testing has been included in to the range of clinical

options for cervical cancer screening. Detection of virus DNA is the basis for HPV testing. HPV testing is known to have some limitations like it is expensive, time-consuming and requires sophisticated laboratory infrastructure. The aim of the study was to evaluate the ability of Raman spectroscopy to detect HPV induced cellular differences in the cell lines. HPV 18 positive HeLa, HPV-16 positive SiHa and HPV negative C33A cell lines were used for the study. The scatter plot showed two clusters belonging to HPV positive and HPV negative cells. Overlap between the clusters of HeLa and SiHa cells was observed. This indicates that cells of these two populations that are HPV expressing cells have similar molecular profile, but showed slight variations. However, these populations of cells (HPV positive) were clustered apart from the HPV non-expressing cells (C33A), exhibiting their differences from HPV negative cells. The findings corroborate with earlier studies that Raman spectroscopy can detect HPV induced cellular effects [118, 119].

Even though HR-HPV detection has clinical significance, as mentioned earlier, it is important to note that very few of HPV-infected subjects eventually develop cancer [142]. Hence, it is important to understand HPV-induced cell changes leading to neoplasia. Further, the differentiation observed in chapter 4, for HPV-positive and -negative cell lines may not be entirely due to HPV presence. This could be because of the fact that the observed spectral variation in HPV-positive and -negative cell lines can be due to differences in cell lines and its origin [4]. This is quite clear from one of Raman Spectroscopy studies on

randomly mixed cell populations [141]. Raman spectral profile very much varies with cell lines and same can be explored for cell typing.

5.4 Future directions

The fundamental goal of optical spectroscopic methods is to provide an objective, non-invasive/minimal-invasive, real-time adjunct for cancer diagnosis/screening. In this study, the feasibility of Raman spectroscopy to classifying normal and abnormal conditions in cervical cancers in a clinical set up was demonstrated. However, further optimization is still desirable, prior to implementation as a routine clinical screening and diagnosis program to prevent cervical cancers. **Specifically, the future work to the thesis study may be directed as follows:**

- ✓ Development of a robust spectroscopy program integrated with comprehensive functional modules, including data acquisition, data process and multivariate statistical analysis is required. To achieve the true real-time diagnosis and characterization, incorporation of a diagnostic model is necessary. Prior to the on-line clinical diagnosis utilizing Raman spectroscopy, a large database must be built to validate the diagnostic models. Optimizing the diagnostic model by employing other multivariate statistic techniques is necessary. In this study, PC-LDA was used to develop diagnostic algorithms throughout the thesis. Apart from PCA and LDA, there exist similar other multivariate statistic techniques needs to be explored 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. To optimize the

diagnosis, a proper selection of multivariate statistic techniques may be one of the choices of method to optimize the diagnostic algorithm. The optimum diagnostic algorithm can also be made user friendly, this may pay its role to clinics.

- ✓ The findings of the study presented in chapter 2 suggest that vaginal sites can be used as internal control, where the normal cervix sites may be unavailable due to advancement of disease. It is important to note that this can also be utilized to bypass inter-patient variability.

- ✓ It is necessary to mimic the live tissue at various histopathological conditions (i.e., normal, benign, LGSILs and HGSILs) for the better understanding of the biochemical changes accounting for Raman *in vivo* 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.

- ✓ The future studies on exfoliated cell specimens may include the development of robust models by selectively accruing abnormal specimens with very higher number of abnormal cells, thus reducing the dominance of normal cells in abnormal specimen in turn and their influence on classification. Once such models are developed, spectra acquired from cell pellet can be compared against model and sample wherein even a single spectrum matches with cancer, it can be assigned as

cancer. This is standard practice in histopathology or cytology. In conventional histopathological examination, several sections are examined and even if one slide show a focus of malignant cells the subjects are treated as cancer.

- ✓ One of the approaches in which standard model can be constructed by utilizing average abnormal spectra from tumor case and average normal spectra from normal case. This approach can be tested by blinded specimens, which may improve the classification efficiency of the model and can avoid the inter-patient variability.

- ✓ Large scale validation Raman study on cervical exfoliated cell specimens needs to be undertaken.

- ✓ In addition to study on cell pellets of exfoliated cervix cells specimens, it's also important to study the biochemical fingerprints of different cell types in these specimens so as to understand the spectral characteristic of a cell pellets.

- ✓ Further studies on HPV-induced-neoplastic changes in the same cell type (i.e same parent cell) are necessary to understand the spectral signatures for these changes.

References

1. Ferlay J S.I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F, Cancer Incidence and Mortality Worldwide: Iarc Cancerbase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer. GLOBOCAN 2012 v1.0, 2013.
2. Lax S., Histopathology of Cervical Precursor Lesions and Cancers. Acta Dermatoven APA, 2011. **20**(3):125-133.
3. Fenn M.B., P. Xanthopoulos, G. Pyrgiotakis, S.R. Grobmyer, P.M. Pardalos and L.L. Hench, Raman Spectroscopy for Clinical Oncology. Advances in Optical Technologies, 2011. **2011**.
4. Rubina S. and C.M. Krishna, Raman Spectroscopy in Cervical Cancers: An Update. J Cancer Res Ther, 2015. **11**(1):10-17.
5. Schiffman M. and N. Wentzensen, Human Papillomavirus Infection and the Multistage Carcinogenesis of Cervical Cancer. Cancer Epidemiology Biomarkers & Prevention, 2013. **22**(4):553-560.
6. Jordan J., A. Singer, H. Jones and M. Shafi, The Cervix2009: John Wiley & Sons.
7. Au W.W., Cervical Cancer – Etiology, Prevention, Intervention and Therapy. Journal of Preventive Medicine, 2005. **13** (3 - 4):9.
8. Viswanathan A.N., M.T. Deavers, A. Jhingran, P.T. Ramirez, C. Levenback and P.J. Eifel, Small Cell Neuroendocrine Carcinoma of the Cervix: Outcome and Patterns of Recurrence. Gynecologic Oncology, 2004. **93**(1):27-33.
9. Denny L., M. Quinn and N. Hacker, Figo Cancer Report 2012. International Journal of Gynaecology and Obstetrics, 2012. **119**(2):8.
10. Sankaranarayanan R., B. Nene, K. Dinshaw, R. Rajkumar, S. Shastri, R. Wesley, P. Basu, R. Sharma, S. Thara and A. Budukh, Early Detection of Cervical Cancer with Visual Inspection Methods: A Summary of Completed and on-Going Studies in India. salud pública de méxico, 2003. **45**:399-407.
11. Kerkar R.A. and Y.V. Kulkarni, Screening for Cervical Cancer: An Overview. J Obstet Gynecol India, 2006. **56**(2):115-122.
12. Solomon D., D. Davey, R. Kurman, A. Moriarty, D. O'Connor, M. Prey, S. Raab, M. Sherman, D. Wilbur and T. Wright Jr, The 2001 Bethesda System: Terminology for Reporting Results of Cervical Cytology. Jama, 2002. **287**(16):2114-2119.
13. Lakowicz J., Principles of Fluorescence Spectroscopy, 1983, 341-381, Plenum Press, New York.
14. Irving J. Bigio S.G.B., Spectroscopic Sensing of Cancer and Cancer Therapy. Cancer Biology and Therapy, 2004. **3**(3):259-267.
15. Ramanujam N., Fluorescence Spectroscopy of Neoplastic and Non-Neoplastic Tissue. . Neoplasia, 2000. **2**(1-2):89-107.
16. Lohmann W., J. Muβmann, C. Lohmann and W. Künzel, 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):249-253.

17. 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):125-127.
18. 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):49-58.
19. 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):647-655.
20. 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):99-107.
21. 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):183-186.
22. 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):46-62.
23. Ramanujam N., M.F. Mitchell, A. Mahadevan, S. Thomsen, A. Malpica, T. Wright, N. Atkinson and R. Richards-Kortum, Spectroscopic Diagnosis of Cervical Intraepithelial Neoplasia (Cin) *in vivo* Using Laser-Induced Fluorescence Spectra at Multiple Excitation Wavelengths. Lasers in surgery and medicine, 1996. **19**(1):63-74.
24. 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, 1994. **91**(21):10193-10197.
25. 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):335-343.
26. Ramanujam N., M.F. Mitchell, A. Mahadevan, S. Thomsen, E. Silva, and R. Richards-Kortum, Rebecca, Fluorescence Spectroscopy: A Diagnostic Tool for Cervical Intraepithelial Neoplasia (Cin). Gynecologic Oncology, 1994. **52**(1):7.
27. Utzinger U., E.V. Trujillo, E. Neely 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. Biomedical Engineering, IEEE Transactions on, 1999. **46**(11):1293-1303.

28. Agrawal A., U. Utzinger, C. Brookner, C. Pitris, M.F. Mitchell and R. Richards-Kortum, Fluorescence Spectroscopy of The. Lasers in surgery and medicine, 1999. **25**:237-249.
29. 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):595-602.
30. 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):1102-1111.
31. 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):67-72.
32. 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):653-658.
33. Atkinson E.N., Age and Fsh Effects in Fluorescence Spectra from the Cervix: An Exploratory Analysis. Gynecologic Oncology, 2005. **99**(3):S95-S97.
34. 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):29-37.
35. 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):523-533.
36. 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):1606-1617.
37. 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. Biomedical Optics, 2007. **12**(3).
38. 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):730-736.

39. 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).
40. Drezek R., I. Boiko, M. Follen, R. Richards-Kortum, A. Malpica, K. Sokolov and U. Utzinger, Understanding the Contributions of Nadh and Collagen to Cervical Tissue Fluorescence Spectra: Modeling, Measurements, and Implications. Journal of Biomedical Optics, 2001. **6**(4):385-396.
41. Pavlova I., K. Sokolov, R. Drezek, A. Malpica, M. Follen, and R. Richards-Kortum, Nadhi Thekkek and Rebecca, Microanatomical and Biochemical Origins of Normal and Precancerous Cervical Autofluorescence Using Laser-Scanning Fluorescence Confocal Microscopy &Para. Photochemistry and photobiology, 2003. **75**(5):550-555.
42. Carlson K., I. Pavlova, T. Collier, M. Descour, M. Follen, and R. Richards-Kortum, Nadhi Thekkek and Rebecca, Confocal Microscopy: Imaging Cervical Precancerous Lesions. Gynecologic Oncology, 2005. **99**(3):S84-S88.
43. 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, NADPH and Collagen as *in vivo* Quantitative Fluorescent Biomarkers of Epithelial Precancerous Changes. Cancer research, 2002. **62**(3):682-687.
44. 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):024008-024008-024012.
45. 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):1223-1236.
46. Milbourne A., S.Y. Park, J.L. Benedet, D. Miller, T. Ehlen, H. Rhodes, A. Malpica, J. Matisic, D. Van Niekirk and E.N. Atkinson, Results of a Pilot Study of Multispectral Digital Colposcopy for the *in vivo* Detection of Cervical Intraepithelial Neoplasia. Gynecologic Oncology, 2005. **99**(3):S67-S75.
47. 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):327-332.
48. 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):947-951.

49. Rodero A.B., L. Silveira Jr, D.A. Rodero, R. Racanicchi and M.T.T. Pacheco, Fluorescence Spectroscopy for Diagnostic Differentiation in Uteri's Cervix Biopsies with Cervical/Vaginal Atypical Cytology. Journal of fluorescence, 2008. **18**(5):979-985.
50. 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):139-145.
51. 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):325-332.
52. 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):587-594.
53. Skala M.C., G.M. Palmer, K.M. Vrotsos, A. Gendron-Fitzpatrick and N. Ramanujam, Comparison of a Physical Model and Principal Component Analysis for the Diagnosis of Epithelial Neoplasias *in vivo* Using Diffuse Reflectance Spectroscopy. Optics Express, 2007. **15**(12):7863-7875.
54. 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):S116-S120.
55. 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):024031-02403111.
56. 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):118-127.
57. 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):064016-064016-064010.
58. Georgakoudi I., B.C. Jacobson, M.G. Müller, E.E. Sheets, K. Badizadegan, D.L. Carr-Locke, C.P. Crum, C.W. Boone, R.R. Dasari and J. Van Dam, Nad (P) H and Collagen as in Vivo Quantitative Fluorescent Biomarkers of Epithelial Precancerous Changes. Cancer research, 2002. **62**(3):682-687.
59. Wong P., 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, 1991. **88**(24):10988-10992.
60. Wong P.T., R.K. Wong and M.F.K. Fung, Pressure-Tuning Ft-Ir Study of Human Cervical Tissues. Applied spectroscopy, 1993. **47**(7):1058-1063.

61. 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):143-153.
62. 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):59-65.
63. Fung M.F.K., M. Senterman, P. Eid, W. Faught, N.Z. Mikhael and P.T. Wong, Comparison of Fourier-Transform Infrared Spectroscopic Screening of Exfoliated Cervical Cells with Standard Papanicolaou Screening. Gynecologic Oncology, 1997. **66**(1):10-15.
64. 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):55-59.
65. 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):47-53.
66. 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):253-257.
67. 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):75-91.
68. 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, 1998. **95**(26):15327-15332.
69. 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):362-366.
70. Romeo M.J., B.R. Wood and D. McNaughton, Observing the Cyclical Changes in Cervical Epithelium Using Infrared Microspectroscopy. Vibrational spectroscopy, 2002. **28**(1):167-175.
71. 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):69-76.
72. 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):577-583.
73. 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):1139-1143.

74. Wood B., L. Chiriboga, H. Yee, M. 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):59-68.
75. 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):145-154.
76. Podshyvalov A., H. Guterman, J. Goldstein, R. Jagannathan, S. Argov, S. Mordechai, S. Mark, R. Sahu and K. Kantarovich, Fourier Transform Infrared Microspectroscopy as a Quantitative Diagnostic Tool for Assignment of Premalignancy Grading in Cervical Neoplasia. Journal of Biomedical Optics, 2004. **9**(3):558-567.
77. Podshyvalov A., R. 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):3725-3734.
78. Raman C., The Colour of the Sea. Nature, 1921. **108**(2716):367-367.
79. Ellis D.I., D.P. Cowcher, L. Ashton, S. O'Hagan and R. Goodacre, Illuminating Disease and Enlightening Biomedicine: Raman Spectroscopy as a Diagnostic Tool. Analyst, 2013. **138**(14):3871-3884.
80. Wang C. and C. Yu, Analytical Characterization Using Surface-Enhanced Raman Scattering (Sers) and Microfluidic Sampling. Nanotechnology, 2015. **26**(9):092001.
81. Colthup N., Introduction to Infrared and Raman Spectroscopy2012: Elsevier.
82. McCreery R.L., Raman Spectroscopy for Chemical Analysis. Vol. 225. 2005: John Wiley & Sons.
83. Carey P., Biochemical Applications of Raman and Resonance Raman Spectroscopes2012: Elsevier.
84. Parker F.S., Applications of Infrared, Raman, and Resonance Raman Spectroscopy in Biochemistry1983: Springer.
85. Das R.S. and Y. Agrawal, Raman Spectroscopy: Recent Advancements, Techniques and Applications. Vibrational spectroscopy, 2011. **57**(2):163-176.
86. Matousek P. and N. Stone, Recent Advances in the Development of Raman Spectroscopy for Deep Non-Invasive Medical Diagnosis. Journal of Biophotonics, 2013. **6**(1):7-19.
87. Jemma G. Kelly J.u.T., Andrew D. Scott, Paul L. Carmichael, Hubert M. Pollock, Pierre L. Martin-Hirsch, and Francis L. Martin, Biospectroscopy to Metabolically Profile Biomolecular Structure: A Multistage Approach Linking Computational Analysis with Biomarkers. J. Proteome Res., 2011. **10**:12.

88. Ruchita S. Das Y.K.A., Raman Spectroscopy: Recent Advancements, Techniques and Applications. Vibrational spectroscopy, 2011. **57** (2011):14.
89. Gabriela Romero E.R., Irina Estrela-Lopis, Edwin Donath and Sergio Enrique Moya, Spontaneous Confocal Raman Microscopy–a Tool to Study the Uptake of Nanoparticles and Carbon Nanotubes into Cells. Nanoscale Research Letters, 2011. **6**(429):4.
90. Liu X., P. Stefanou, B. Wang, T. Woggon, T. Mappes and U. Lemmer, Organic Semiconductor Distributed Feedback (Dfb) Laser as Excitation Source in Raman Spectroscopy. Optics Express, 2013. **21**(23):28941-28947.
91. Liang R., Biomedical Optical Imaging Technologies: Design and Applications2012: Springer.
92. Chatfield C. and A. Collins, Introduction to Multivariate Analysis. Vol. 1. 1981: CRC Press.
93. Tietz N. and R. Dryer, Fundamentals of Clinical Biochemistry. Saunders, New York, 1976.
94. Chatfield C., The Analysis of Time Series: An Introduction2013: CRC press.
95. Manly B.F., Multivariate Statistical Methods Chapman & Hall. New York, New York, USA, 1994.
96. Ellis D.I. and R. Goodacre, Metabolic Fingerprinting in Disease Diagnosis: Biomedical Applications of Infrared and Raman Spectroscopy. Analyst, 2006. **131**(8):875-885.
97. Polat K. and S. Güneş, Principles Component Analysis, Fuzzy Weighting Pre-Processing and Artificial Immune Recognition System Based Diagnostic System for Diagnosis of Lung Cancer. Expert Systems with Applications, 2008. **34**(1):214-221.
98. Plsplus/Iq User's Guide, . Galactic Industries Corporation, Salem, New Hampshire, 1999.
99. Devore J., Probability and Statistics for Engineering and the Sciences2011: Cengage Learning.
100. Teh S., W. Zheng, K. Ho, M. Teh, K. 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):457-465.
101. Ghanate A., S. Kothiwale, S. Singh, D. Bertrand and C.M. Krishna, Comparative Evaluation of Spectroscopic Models Using Different Multivariate Statistical Tools in a Multicancer Scenario. Journal of Biomedical Optics, 2011. **16**(2):025003-025003-025009.
102. Zeng H. Real-Time *in vivo* Raman Spectroscopy-Technology Development and Clinical Applications in Early Cancer Detection. in *Bio-Optics: Design and Application*. 2013. Optical Society of America.
103. Tu Q. and C. Chang, Diagnostic Applications of Raman Spectroscopy. Nanomedicine: Nanotechnology, Biology and Medicine, 2012. **8**(5):545-558.

104. Crow P., A. Molckovsky, N. Stone, J. Uff, B. Wilson and L.-M. Wongkeesong, Assessment of Fiberoptic near-Infrared Raman Spectroscopy for Diagnosis of Bladder and Prostate Cancer. Urology, 2005. **65**(6):1126-1130.
105. 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):123-132.
106. 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):427-431.
107. 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**(8):955-959.
108. Robichaux-Viehoever A., E. Kanter, H. Shappell, D. Billheimer, H. Jones III and A. Mahadevan-Jansen, Characterization of Raman Spectra Measured *in vivo* for the Detection of Cervical Dysplasia. Applied spectroscopy, 2007. **61**(9):986-993.
109. Kanter E.M., E. Vargis, S. Majumder, M.D. Keller, E. Woeste, G.G. Rao and A. Mahadevan-Jansen, Application of Raman Spectroscopy for Cervical Dysplasia Diagnosis. Journal of Biophotonics, 2009. **2**(1-2):81-90.
110. 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):205-211.
111. 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):512. e511-512. e515.
112. Mo J., W. Zheng, J.J. Low, J. Ng, A. Ilancheran and Z. Huang, High Wavenumber Raman Spectroscopy for *in vivo* Detection of Cervical Dysplasia. Analytical chemistry, 2009. **81**(21):8908-8915.
113. Duraipandian S., W. Zheng, J. Ng, J.J. Low, A. Ilancheran and Z. Huang, *In Vivo* Diagnosis of Cervical Precancer Using Raman Spectroscopy and Genetic Algorithm Techniques. Analyst, 2011. **136**(20):4328-4336.
114. Vargis E., T. Byrd, Q. Logan, D. Khabele and A. Mahadevan-Jansen, Sensitivity of Raman Spectroscopy to Normal Patient Variability. Journal of Biomedical Optics, 2011. **16**(11):117004-1170049.
115. Duraipandian S., W. Zheng, J. Ng, J.J. Low, A. Ilancheran and Z. Huang, Simultaneous Fingerprint and High-Wavenumber Confocal Raman Spectroscopy Enhances Early Detection of Cervical Precancer *in vivo*. Analytical chemistry, 2012. **84**(14):5913-5919.

116. Duraipandian S., W. Zheng, J. Ng, J.J. Low, A. Ilancheran and Z. Huang, Near-Infrared-Excited Confocal Raman Spectroscopy Advances *in vivo* Diagnosis of Cervical Precancer. Journal of Biomedical Optics, 2013. **18**(6):067007-067007.
117. Jess P.R., D.D. 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):2723-2728.
118. Ostrowska K.M., A. Malkin, A. Meade, J. O'Leary, C. Martin, C. Spillane, H.J. Byrne and F.M. Lyng, Investigation of the Influence of High-Risk Human Papillomavirus on the Biochemical Composition of Cervical Cancer Cells Using Vibrational Spectroscopy. Analyst, 2010. **135**(12):3087-3093.
119. Vargis E., Y.-W. Tang, D. Khabele and A. Mahadevan-Jansen, Near-Infrared Raman Microspectroscopy Detects High-Risk Human Papillomaviruses. Translational oncology, 2012. **5**(3):172.
120. Krishna C.M., N. Prathima, R. Malini, B. Vadhiraja, R.A. Bhatt, D.J. Fernandes, P. Kushtagi, M. Vidyasagar and V. Kartha, Raman Spectroscopy Studies for Diagnosis of Cancers in Human Uterine Cervix. Vibrational spectroscopy, 2006. **41**(1):136-141.
121. Downes A. and A. Elfick, Raman Spectroscopy and Related Techniques in Biomedicine. Sensors, 2010. **10**(3):1871-1889.
122. Faolain E.O., M.B. Hunter, J.M. Byrne, P. Kelehan, H.J. Byrne and F.M. Lyng. The Potential of Vibrational Spectroscopy in the Early Detection of Cervical Cancer: An Exciting Emerging Field. in *OPTO-Ireland*. 2005. International Society for Optics and Photonics.
123. Lyng F.M., E.Ó. Faoláin, J. Conroy, A. Meade, P. Knief, B. Duffy, M. Hunter, J. Byrne, P. Kelehan and H.J. Byrne, Vibrational Spectroscopy for Cervical Cancer Pathology, from Biochemical Analysis to Diagnostic Tool. Experimental and molecular pathology, 2007. **82**(2):121-129.
124. da Silva Martinho H., M.C.B.M. Yassoyama, P. de Oliveira Andrade, R.A. Bitar, A.M. do Espírito Santo, E.Â.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):054029-054029-054026.
125. 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):255-261.
126. Wachsmann-Hogiu S., T. Weeks and T. Huser, Chemical Analysis *in vivo* and *in vitro* by Raman Spectroscopy—from Single Cells to Humans. Current opinion in biotechnology, 2009. **20**(1):63-73.
127. Gobinet C., V. Vrabie, M. Manfait and O. Piot, Preprocessing Methods of Raman Spectra for Source Extraction on Biomedical Samples: Application on Paraffin-Embedded Skin Biopsies. Biomedical Engineering, IEEE Transactions on, 2009. **56**(5):1371-1382.

128. Nijssen A., L.F. Santos, P.J. Caspers, J.C. den Hollander, M.H. Neumann, G.J. Puppels, T.C.B. Schut and K. Maquelin, Discriminating Basal Cell Carcinoma from Perilesional Skin Using High Wave-Number Raman Spectroscopy. Journal of Biomedical Optics, 2007. **12**(3):034004-034004-034007.
129. Koljenović S., T. Lin-P'ing Choo-Smith, C.B. 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):1265-1277.
130. Singh S., A. Sahu, A. Deshmukh, P. Chaturvedi and C.M. Krishna, In Vivo Raman Spectroscopy of Oral Buccal Mucosa: A Study on Malignancy Associated Changes (Mac)/Cancer Field Effects (Cfe). Analyst, 2013. **138**(14):4175-4182.
131. Young B., P. Woodford and G. O'Dowd, Wheater's Functional Histology: A Text and Colour Atlas2013: Elsevier Health Sciences.
132. Sankaranarayanan R., B.M. Nene, S.S. Shastri, K. Jayant, R. Muwonge, A.M. Budukh, S. Hingmire, S.G. Malvi, R. Thorat and A. Kothari, Hpv Screening for Cervical Cancer in Rural India. New England Journal of Medicine, 2009. **360**(14):1385-1394.
133. Walsh M.J., M.N. Singh, H.M. Pollock, L.J. Cooper, M.J. German, H.F. Stringfellow, N.J. Fullwood, E. Paraskevaidis, P.L. Martin-Hirsch and F.L. Martin, Atr Microspectroscopy with Multivariate Analysis Segregates Grades of Exfoliative Cervical Cytology. Biochemical and biophysical research communications, 2007. **352**(1):213-219.
134. Jamie M.N., Fixation and Tissue Processing. Special Stains and H & E, ed. G.L. Kumar J.A.K.2010, Carpinteria, North America,: Dako. 141–152.
135. Zare-Mirzaie A. and K.K.-A.M. Abolhasani, Rehydration of Air-Dried Cervical Smears: An Alternative to Routine Wet Fixation. Acta Medica Iranica, 2007. **45**(5):365-368.
136. Deodhar K., R. Sankaranarayanan, K. Jayant, J. Jeronimo, R. Thorat, S. Hingmire, R. Muwonge, A. Chiwate, R. Deshpande and D. Ajit, Accuracy of Concurrent Visual and Cytology Screening in Detecting Cervical Cancer Precursors in Rural India. International journal of cancer, 2012. **131**(6):E954-E962.
137. Virkler K. and I.K. Lednev, Raman Spectroscopic Signature of Blood and Its Potential Application to Forensic Body Fluid Identification. Analytical and bioanalytical chemistry, 2010. **396**(1):525-534.
138. Pachmann K., J.H. Clement, C.-P. Schneider, B. Willen, O. Camara, U. Pachmann and K. Höffken, Standardized Quantification of Circulating Peripheral Tumor Cells from Lung and Breast Cancer. Clinical Chemical Laboratory Medicine, 2005. **43**(6):617-627.
139. Loos J. and D. Roos, Ficoll-Isopaque Gradients for the Determination of Density Distributions of Human Blood Lymphocytes and Other Reticulo-Endothelial Cells. Experimental cell research, 1974. **86**(2):333-341.
140. Sato E., S.H. Olson, J. Ahn, B. Bundy, H. Nishikawa, F. Qian, A.A. Jungbluth, D. Frosina, S. Gnjatic and C. Ambrosone, Intraepithelial Cd8+

Tumor-Infiltrating Lymphocytes and a High Cd8+/Regulatory T Cell Ratio Are Associated with Favorable Prognosis in Ovarian Cancer. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(51):18538-18543.

141. C. Murali Krishna G.D.S., Gregory Kegelaera, Sylvain Rubina, Vasudevan B. Karthab, Michel Manfaita, Micro-Raman Spectroscopy of Mixed Cancer Cell Populations. Vibrational spectroscopy, 2004. **38**(1-2):6.
142. Chen H.-C., M. Schiffman, C.-Y. Lin, M.-H. Pan, S.-L. You, L.-C. Chuang, C.-Y. Hsieh, K.-L. Liaw, A.W. Hsing and C.-J. Chen, Persistence of Type-Specific Human Papillomavirus Infection and Increased Long-Term Risk of Cervical Cancer. Journal of the National Cancer Institute, 2011. **103**(18):1387-1396.

Appendix I: Publications and reprints