Laser Raman Spectroscopy as an Adjunct \textit{in vivo} Endoscopic Device for Improving Early Lung Cancer Detection

by

Hanna Claire McGregor

B.Sc., The University of British Columbia, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Interdisciplinary Oncology)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2018

© Hanna Claire McGregor, 2018
Abstract

Lung cancer has an 18% five year survival rate, which is among the lowest in all cancer types. Typically, this low survival rate is due to the detection of the disease at a late stage not amendable to curative therapy. To combat this poor prognosis, many efforts have been made to detect early lung cancers before they become metastatic. Diagnosis of lung cancer involves localization and biopsy but the diagnostic accuracy of small lung lesions is currently sub-optimal. In the central airways autofluorescent bronchoscopy can localize suspicious areas for biopsy with high sensitivity but the specificity is relatively low. In the peripheral airways abnormal tissue is localized through the radial endobronchial ultrasound procedure. The diagnostic yield is relatively low in the sixty percent range, therefore, there remains a need for real time detailed information about benign versus malignant lung tissue for endoscopic diagnosis.

Raman spectroscopy, a technique that utilizes the inelastic scattering of light by a molecule, has the ability to show detailed biochemical information that current procedures do not provide. Here we study the feasibility of incorporating Raman spectroscopy into standard clinical procedures. In the central airways we conduct a large scale single centre trial using Raman Spectroscopy as an adjunct to autofluorescence bronchoscopy procedures. It was found that adjunct RS can greatly improve the classification of malignant from benign/normal lesions with a high diagnostic sensitivity (90%) and good specificity (65%). In the peripheral airways, we present the design and development of a novel miniature Raman probe capable of navigating the bronchial architecture into the small airways. To our knowledge, we show the first in vivo clinical test of a fibre bundle peripheral probe and the first Raman spectra of peripheral lung
cancer and normal tissue. With follow up clinical testing and validation we believe the opportunity to use Raman spectroscopy as an adjunct device in the entire lung is feasible.
Lay Summary

The five year lung cancer survival rate in Canada remains low at only 18%. While emerging detection, screening and localization strategies are working to improve this survival rate, many technologies currently in use have a high false positive rate, low specificity or low diagnostic yield. There is a need for real time in vivo diagnosis strategies at the bedside that can improve on the limitations of current technologies. Here we present a large-scale single center investigation which showed that endoscopic point laser Raman spectroscopy was able to differentiate malignant central lung lesions from benign/normal. Further, we discuss the design and development of a novel lung probe capable of navigating into the peripheral airways and we present the first ever clinical in vivo test of a Raman device used in the small airways.
Preface

Partial content of Chapter 1 has been published: McGregor, H.C., Wang, W., Short, M.A., and Zeng, H. Clinical utility of Raman spectroscopy: current applications and ongoing developments. (2016). Advanced Heath Care Technologies. Invited Review. For this publication McGregor H.C. conducted the literature search, review and manuscript preparation. Wang, W. provided assistance and comments with manuscript preparation. McGregor H.C., and Wang, W., designed, filmed, and edited the video abstract. Short M.A., and Zeng, H provided comments to the manuscript. This publication is ref. (1) throughout this thesis.

Partial content of Chapter 2 has been published: Pawluk H.C., Short, M.A., Lam, S., McWilliams, A.M., Ionescu, D.N., Zeng, H. Improvements to a laser Raman spectroscopy system for reducing the false positives of autofluorescence bronchoscopy. Proc. SPIE 8207, Photonic Therapeutics and Diagnostics VIII, 820737 (Feb 2012). For this publication Zeng, H. and Short, M.A. identified and designed the study. Pawluk (now McGregor) H.C., performed all the experiments, analyzed the data and prepared the manuscript. Short M.A., Lam, S., McWilliams A.M., and Ionescu D.N., Zeng H. provided comments to the manuscript. This publication is ref. (2) throughout this thesis.

collected and analyzed the macro-Raman data. Hewitt K.C., and Brouwers E., prepared the manuscript and all authors provided comments to the manuscript. This publication is ref. (3) throughout this thesis and can be viewed at:

http://pubs.rsc.org/en/content/articlehtml/2015/an/c5an01080b

Partial content of Chapter 2 has been published: Sonis S., Guze, K., Pawluk H.C., Short, M.A., Zeng, H., Lorch, J., and Norris, C. A Pilot Study that shows that Raman Spectroscopy effectively differentiates premalignant and malignant lesions from normal mucosa and benign lesions in humans. Head and Neck, 2015, 37, 511-517. For this publication Guze, K., Zeng, H., Norris, C., and Sonis, S., designed the study. Guze, K., collected the data and prepared the manuscript. Pawluk (now McGregor), H. C., and Short, M. A., analyzed the data and helped prepare the manuscript. All authors provided comments to the manuscript. This publication is ref. (4) throughout this thesis.

Content of Chapter 3 has been published: McGregor, H. C., Short, M. A., McWilliams, A.M., Shaipanich, T., Ionescu, D. N., Zhao, J., Wang, W., Chen, G., Lam, S. and Zeng, H. (2016), Real-time endoscopic Raman spectroscopy for in vivo early lung cancer detection. J. Biophoton, 10: 98–110. For this publication Zeng, H., Short, M.A., and Lam, S., designed the study. Lam, S., McWilliams, A.M., and Shaipanich, T., performed the clinical bronchoscopy procedures. Ionescu, D. N., determined the pathology of biopsy specimens. McGregor H.C., performed the in vivo measurements, conducted data analysis and prepared the manuscript. Zhao, J., helped with data analysis and manuscript preparation. Short, M.A., Wang, W., and Chen, G., provided insight into data analysis. All authors provided comments on the manuscript. This publication is ref. (5) throughout this thesis.
Content of Chapter 4 and 5 has been submitted for publication: McGregor, H.C., Short, M.A., Lam, S., Shaipanich, T., Beaudoin, E-L., and Zeng, H. Development and in vivo testing of a Raman spectroscopy system for the detection of peripheral lung lesions. For this publication Zeng, H., McGregor, H.C., Short, M. A., and Lam, S. designed the study. McGregor, H.C. and Short, M. A., designed and developed the novel lung probe. McGregor, H.C., performed all experiments, analyzed the data and prepared the manuscript. Lam, S., Shaipanich, T., and Beaudoin, E-L., performed the clinical bronchoscopy procedures. All authors provided comments on the manuscript.

Ethics approval for the collection of Raman spectral data from lung patients was obtained from the University of British Columbia – BC Cancer Agency Research Ethics Board (REB #: H06-00010).
Table of Contents

Abstract ................................................................................................................................. ii

Lay Summary ........................................................................................................................ iv

Preface .................................................................................................................................. v

Table of Contents ................................................................................................................... viii

List of Tables .......................................................................................................................... xiii

List of Figures ........................................................................................................................ xiv

List of Abbreviations .......................................................................................................... xviii

Acknowledgements .............................................................................................................. xxii

Dedication ............................................................................................................................. xxiv

Chapter 1: Introduction ........................................................................................................ 1

1.1 Introduction to Lung Cancer .......................................................................................... 1

1.2 Lung Cancer Progression ............................................................................................. 2

1.2.1 Lung Cancer in Central Airways ........................................................................... 2

1.2.2 Lung Cancer in Peripheral Airways ....................................................................... 4

1.3 Current Clinical Diagnostic Methods ......................................................................... 6

1.3.1 Detection Methods ................................................................................................. 7

1.3.2 Localization and Biopsy Retrieval Methods .......................................................... 8

1.4 Clinical Gaps ................................................................................................................ 12
1.4.1 Central Airways ................................................................. 12
1.4.2 Peripheral Airways ........................................................... 12
1.5 Raman Spectroscopy ............................................................. 13
1.6 Thesis Aims and Outline ....................................................... 18

Chapter 2: Optimization of Hardware and Data Analysis Methods ...... 20
2.1 Chapter Introduction ............................................................ 20
2.2 Instrumentation ................................................................. 20
2.2.1 Old versus New System Design ......................................... 21
2.2.2 Standard Comparison ....................................................... 23
2.3 Spectral Pre-processing Techniques and Multivariate Statistical Methods .......... 26
2.3.1 Spectral Pre-processing Techniques ...................................... 26
2.3.2 Multivariate Statistical Methods ......................................... 29
2.4 Performance of Pre-processing and Analytical Techniques using \textit{ex vivo} Data Sets ........ 35
2.4.1 \textit{Ex Vivo} Liver Data Sets ............................................... 35
2.5 Development and Validation of \textit{in vivo} Analytical Techniques ......................... 40
2.5.1 \textit{In Vivo} Oral Cavity Data Set ........................................... 40
2.6 Chapter Conclusions ........................................................... 47

Chapter 3: Real-time endoscopic Raman spectroscopy for \textit{in vivo} early lung cancer detection in the Central Airways ................................................................. 49
3.1 Chapter Introduction ............................................................ 49
3.2 Patients and Methods .................................................................................................. 50
3.3 Statistical Analysis ................................................................................................. 57
3.4 Results and Discussion ........................................................................................ 61
  3.4.1 Mean Raman Spectra Peak Assignments ......................................................... 61
  3.4.2 Multivariate Statistical Analysis ..................................................................... 64
  3.4.3 Sensitivity and Specificity of Visual Diagnosis .............................................. 77
3.5 Chapter Conclusions ............................................................................................. 81

Chapter 4: Design and Development of an in vivo Peripheral Lung Raman Catheter ...... 82
  4.1 Chapter Introduction ............................................................................................. 82
  4.2 Challenges for Peripheral Devices and Design Rationale ................................. 83
  4.3 Developmental Procedure ................................................................................ 85
    4.3.1 Ferrule Removal ............................................................................................ 87
    4.3.2 Etching Procedure ...................................................................................... 88
    4.3.3 Epoxy Procedure ......................................................................................... 90
    4.3.4 Final Peripheral Probe Design ................................................................. 92
  4.4 Measurements on Tylenol and in vivo Skin ....................................................... 94
  4.5 Chapter Conclusions .......................................................................................... 98

Chapter 5: Proof of principle in vivo testing of a Raman spectroscopy system for the
detection of peripheral lung lesions ........................................................................... 99
  5.1 Chapter Introduction ............................................................................................ 99
Chapter 6: Case study: in vivo Raman spectrum of an Aspergilloma ........................................... 111

6.1 Chapter Introduction .................................................................................................................. 111

6.2 Materials and Methods .......................................................................................................... 112

6.2.1 In Vivo Measurement ......................................................................................................... 112

6.2.2 In Vitro Fungal Measurements ............................................................................................ 113

6.3 Results and Discussion ......................................................................................................... 117

6.3.1 In Vivo Measurement with the Peripheral Lung Raman System ........................................ 117

6.3.2 In Vitro Measurements with the Peripheral Lung Raman System ................................. 119

6.3.4 Spectral Comparison ......................................................................................................... 121

6.4 Chapter Conclusion .............................................................................................................. 123

Chapter 7: Conclusions and Future Directions ......................................................................... 124

7.1 Conclusions .......................................................................................................................... 124

7.2 Current Limitations and Future Directions ........................................................................... 126

Bibliography .................................................................................................................................. 128

Appendices ..................................................................................................................................... 135

Appendix A ..................................................................................................................................... 135

Appendix B ..................................................................................................................................... 139
List of Tables

Table 2.1: Summary of samples measured and LDA classification success constrained by the requirement of all lesions with a pathology ≥MILD have to be identified.................. 47

Table 3.1: Physician visual grading system and the corresponding tissue description.................. 55

Table 3.2: Pathological coding system and the corresponding tissue diagnosis.......................... 56

Table 3.3: Patient demographics and location of the lung Raman readings............................ 57

Table 3.4: Area under the ROC curve and 95% CI based on full range and waveband selection algorithms for discriminating premalignant (HGD) and malignant lung cancers (n = 72) from benign lung diseases and normal lung tissues (n = 208)............................ 66

Table 3.5: Summary of Raman spectroscopy diagnostic parameters derived from ROCs........... 67

Table A.1: TNM Staging of Lung Cancer .................................................................................. 135

Table A.2: Structural Anatomy of Each Airway Generations ................................................. 137

Table A.3: Pathology Details for Premalignant Squamous Cell Central Airway Lesions ...... 138
List of Figures

Figure 1.1: Schematic showing the difference between Rayleigh, Stokes and Anti-Stokes Raman Scattering ................................................................. 15

Figure 2.1: Schematic diagram of the new in vivo Raman system (not to scale), with magnified views of the fiber tips ................................................................. 23

Figure 2.2: The average of 100 Raman Spectra of Tylenol using both the new and old system 24

Figure 2.3: The signal to noise ratio of both the old system and the new system ....................... 25

Figure 2.4: The fluorescence background removal using the Vancouver Algorithm .................. 28

Figure 2.5: Overview of the preprocessing and statistical techniques used throughout this thesis ........................................................................................................... 35

Figure 2.6: Average Raman spectra of normal and fatty intact mice livers using the newly updated Raman system ................................................................. 36

Figure 2.7: Difference spectra generated by subtraction of the control (Normal) mice spectra from the fatty mice spectra ................................................................. 37

Figure 2.8: The contribution of factor one converted back into a spectrum for each week on the fatty liver diet ............................................................................... 39

Figure 2.9: Averaged ≥MILD (orange) and averaged ≤LEU (purple) Raman spectra after normalization by area under the curve ......................................................... 43

Figure 2.10: Posterior Probability plot of an oral lesion being classified of ≥MILD .................. 45

Figure 2.11: Receiver Operator Characteristic (ROC) ................................................................................. 46

Figure 3.1: Schematic diagram of the endoscopic laser Raman Spectroscopy system ............... 51

Figure 3.2: An example showing 9th order polynomial fitting to the fluorescence background of the raw spectra, and the resultant Raman spectra ......................................................... 53
Figure 3.3: Mean Raman spectra by diagnosis. ................................................................. 62

Figure 3.4: Mean Raman spectra by diagnosis, with highlighted areas showing the selected wavebands by LASSO ................................................................. 69

Figure 3.5: Mean Raman spectra by diagnosis, with highlighted areas showing the selected wavebands by Stepwise ................................................................. 70

Figure 3.6: Lesion classification by Raman spectroscopy based on STEP PC-GDA analysis.. 71

Figure 3.7: Lesion classification by Raman spectroscopy based on STEP PC-GDA analysis (box plot) ........................................................................................................ 72

Figure 3.8: Lesion classification based on modeling using the extremes cases (29 Invasive Cancers and 118 Normal Sites) for training (full spectra PC-GDA) ............... 75

Figure 3.9: Lesion classification based on modeling using the extremes cases (29 Invasive Cancers and 118 Normal Sites) for training (full spectra PC-GDA) (box plot) ..... 76

Figure 3.10: The histopathology distribution of bronchoscopy (WLB+AFB) visual grading of lung lesions and normal tissue sites ......................................................... 78

Figure 4.1: Specifications for the first generation peripheral lung Raman Probe .................. 86

Figure 4.2: End on view of the distal end of the first generation peripheral lung Raman probe 87

Figure 4.3: Side view of the distal end of the peripheral lung Raman probe (prior to the metal ferrule removal). ............................................................................................... 89

Figure 4.4: Final design of the peripheral lung Raman probe ........................................... 93

Figure 4.5: End on view (A) and side view (B) of the final functional peripheral lung Raman probe ........................................................................................................... 94

Figure 4.6: The mean spectra (n=100) of Tylenol collected with the newly developed peripheral lung Raman probe. ............................................................................... 95
Figure 4.7: The mean Raman spectra (n=100) of Tylenol collected with the central lung probe (bottom) and the peripheral lung probe (top) ........................................... 96

Figure 4.8: The mean (n=100) Raman spectrum of in vivo palm skin of a 28 year old Caucasian volunteer measured using the newly developed peripheral probe. ....................... 97

Figure 5.1: The standard clinical procedure for a peripheral lesion needing a biopsy (left) and the Raman procedure (right). ................................................................. 102

Figure 5.2: The clinical procedure for locating and measuring peripheral lesions with the miniature Raman catheter ................................................................. 103

Figure 5.3: Size (24.2 mm by 22.9 mm) and location (RB1) of a cancerous peripheral nodule found using EBUS. ................................................................. 104

Figure 5.4: Mean normalized spectra of each pathology type. Normal (n = 31), Cancer (n=19) and Whole Blood (n=23). Spectra are separated on the y-axis for clarity......... 107

Figure 5.5: Mean differences in Raman spectra from peripheral tissue with regions of significant difference (p-value ≤ 0.005) superimposed. Spectra are separated on the y-axis for clarity................................................................. 109

Figure 6.1: The mechanical design of the airtight optic al container base used to hold the A. fumigatus samples................................................................. 114

Figure 6.2: The mechanical design of the airtight optical container cap used to hold the A. fumigatus samples................................................................. 115

Figure 6.3: Aspergillus fumigatus samples loaded into specially designed airtight optical containers ................................................................. 116

Figure 6.4: The in vivo mean (n=13) spectra of an aspergilloma................................................................. 118
Figure 6.5: The mean (n=100) *in vitro* spectra of the four *Aspergillus fumigatus* samples acquired with the PLRS ................................................................. 120

Figure 6.6: Mean spectra using the PLRS for each media type .......................................... 121

Figure 6.7: The mean spectra of the *in vivo* Aspergilloma measurement (n=13), the *in vivo* normal lung tissue measurement (n=15) and the *in vitro* fungal sample measurement using the PLRS (n=400) ........................................................................... 122

Figure B.1: An example raw spectrum collected from *ex vivo* Rat Liver tissue. .............. 139

Figure B.2: Example calculation showing the SG smoothing function over 3 different window sizes. ................................................................................................................. 142

Figure B.3: An example Smoothed (via SG) Raw spectrum collected from *ex vivo* Rat Liver tissue. ................................................................................................................. 143

Figure B.4: Example calculation showing the SG 2nd order derivative function over an 11 point window size ............................................................................................................. 144

Figure B.5: An example 2nd order derivative spectrum collected from *ex vivo* Rat Liver tissue. ................................................................................................................. 145

Figure B.6: Example calculation showing the normalization procedure of the SG 2nd order derivatives................................................................................................................. 146

Figure B.7: An example of a normalized 2nd order derivative spectrum collected from *ex vivo* Rat Liver tissue. ............................................................................................................. 147
List of Abbreviations

AAH: Atypical Adenomatous Hyperplasia

ADC: Adenocarcinoma

AF: Autofluorescence

AFB: Autofluorescent Bronchoscopy

AIS: Adenocarcinoma \textit{in situ}

ANSI: American National Standards Institute

AUC: Area Under the Curve

AURA: Aura Skin System

BCCA VC: British Columbia Cancer Agency Vancouver Centre

CCD: Charge Coupled Device

CIS: Carcinoma \textit{in situ}

CT: Computed Tomography

CT-TTLB: CT guided Trans-thoracic Lung Biopsy

DA: Discrimination Analysis

DNA: Deoxyribonucleic Acid

FEP: Fluorinated Ethylene Propylene

GA: Genetic Algorithm
GDA: Generalized Discrimination Analysis

HGD: High Grade Dysplasia

IC: Invasive Carcinoma

LASSO: Lease Absolute Shrinkage and Selection Operator

LDA: Linear Discrimination Analysis

LDCT: Low Dose Computed Tomography

≤LEU: Leukoplakia and Better

LLL: Left Lower Lung

LOO-CV: Leave One Out Cross Validation

LUL: Left Upper Lung

MIA: Minimally Invasive Adenocarcinoma

≥MILD: Mild Dysplasia and Worse

MWU: Mann Whitney U

NIPALS: Nonlinear Iterative Partial Least Squares

NLST: National Lung Cancer Screening Trial

NSCLC: Non-Small Cell Lung Cancer

PC: Personal Computer
PCA: Principle Component Analysis

PCs: Principle Components

PLRS: Peripheral Lung Raman System

PLS: Partial Least Squares

R-EBUS: Radial Endobronchial Ultrasound

R/G: Red/Green Ratio

RLL: Right Lower Lobe

RML: Right Middle Lobe

ROC: Receiver Operator Characteristic

ROS: Reactive Oxygen Species

RPMI: Roswell Park Memorial Institute Medium

RS: Raman Spectroscopy

RUL: Right Upper Lobe

SCC: Squamous Cell Carcinoma

SCLC: Small Cell Lung Cancer

SNR: Signal to Noise Ratio

STEP: Forward Stepwise
TMN: Tumor Metastasis Nodes

VA: Vancouver Algorithm

WHO: World Health Organization

WLB: White Light Bronchoscopy

YAG: Yeast Agar Glucose media
Acknowledgements

I would like to express my sincere appreciation for my supervisor, Dr. Haishan Zeng, for whom this work would not be possible without. I thank Dr. Zeng for the unwavering support, guidance and commitment to my learning that he has shown over these years. I thank him for the time spent on reviewing scholarship applications, manuscript drafts, and power-point presentations, all while teaching me how to be an independent researcher. The insight Dr. Zeng has provided over the course of my graduate education is unparalleled.

I thank my dedicated lab members, who had an openness and willingness to teach me the aspects of engineering that I greatly needed. In particular, I would like to thank Dr. Michael Short, who was willing to teach me for years: everything from the general concepts of Raman spectroscopy, to the processes of medical device development, to the finite details of mathematical modelling. Dr. Short has been an integral part of my growth as a scientist and as such, will have my lasting gratitude. As well, thank you to Dr. Jianhua Zhao for the technical explanations and assistance navigating various programs like MatLAB and LABView, and for the advice on interpreting spectra and various statistical results.

Thank you to my committee members, Dr. Calum MacAulay and Dr. Stephen Lam, for guiding me throughout the scientific process, providing much needed expertise, and pushing me to ask hard questions. Thank you as well to the physician team, Dr. Tawimas Shaipanich, Dr. Annette McWilliams, Dr. Eve-Lea Beaudoin and the clinical team at the BCCA Vancouver Centre. The surgical daycare team has gone beyond the call of duty to help make this in vivo work possible; especially Myles McKinnon, who I’d like to recognize for his insight into surgical procedure, willingness to coordinate cases, and well timed humor.
I wish to thank Dr. Anthony Lee for guidance and encouragement during the final months of data collection and thesis writing; and for teaching me the value of lunch. Thank you to Geoffrey Hohert for the technical expertise and patience to teach me (often multiple times) the operations of the 3D printer, polisher, and SolidWorks. Thank you to my current and (former) lab mates: Giselle Tian, Zhenguow Wu, (Tracey Wang), (Kam Chow), (Wenbo Wang) and Maryam Shirmohammad who were responsible for making this PhD experience unforgettable.

I wish to thank my Dad, Gary Pawluk, who regularly put his own life on hold in order to see me succeed. For the sacrifice of time, money and sometimes sanity, I owe him special recognition. He is constantly teaching me life skills outside the scope of academia with the patience only a father could have. For personifying and exemplifying the word ‘tenacity’, being my driven inner voice, and being the best card partner I’ve ever had, my appreciation is endless.

As this thesis is for the Interdisciplinary Oncology Program, I would be remiss if I did not acknowledge my Mom, Tracey Pawluk. I thank her for introducing me into the field when I was 7, for helping me apply to universities when I was 17, for driving me to the BCCRC for the first time when I was 21, for the reassurance before my comprehensive exam at 23, and for the heartfelt relief when I was – finally – graduated at 28. Yet more importantly, for showing me every day in between that the word ‘survivor’ means much more than a statistical grouping. For that I owe my mother gratitude far beyond the limits of these pages.

Finally, thank you to my husband Justin McGregor. From years of paying tuition together, to wine weekends away, to finding Lily, to all the late nights, hectic timetables, and emotional turbulence: Justin has supported me in the highs, comforted me in the lows, and in between provided a stability to our family that I am so grateful for. To quote TC, Justin is not the husband I deserve but the husband that I need.
Dedication

For Mom, Dad and Justin
Chapter 1: Introduction

1.1 Introduction to Lung Cancer

Lung cancer is the leading cause of cancer-related deaths worldwide (6). In Canada during 2016 more than 20,000 deaths were due to lung cancer alone (7). Although smoking related lung cancer accounts for the vast majority of diagnosed cases, non-smoking related lung cancer is still estimated to account for 10-15% of all lung cancer cases (8).

The overall five year survival rate of lung cancer is 18% after diagnosis (6). The most significant reason for the poor prognosis is late stage detection, where there is no effective curative treatment. This is in contrast to patients found with early stage (0) carcinoma in situ (CIS) or stage 1A (tumor less than 2 cm without metastatic spread) where the 5 year survival is >80% (9). The WHO TMN staging criteria, shown in Appendix A Table A.1, is used in lung cancer management (10, 11). Similar to other cancer types, detection of lung cancer at early stages leads to a better prognosis for the patient.

Lung cancer encompasses many different cancer subtypes. Classification of lung cancer is first divided into two major groups, Small-Cell Lung Cancer (SCLC) and Non-Small Cell Lung Cancer (NSCLC), with NSCLC making up 85% of all lung cancer cases (12). The NSCLC group can be divided further into three more subtypes which included Squamous Cell Carcinoma (SCC), Adenocarcinoma (ADC) and Large Cell Carcinoma.

Treatment strategies for lung cancer depend on the stage of the disease, tumor subtype, the patient’s overall lung health and whether the disease can be resected. Surgery is the primary choice of treatment if the disease has not metastasized and the patient is fit enough for surgery and life with reduced lung capacity (13). Radiation therapy and chemotherapy are secondary modes of treatment for patients with disease that has invaded local lymph nodes or spread.
outside the chest. (13). Immunotherapy is used for metastatic cases, or tumors which are no longer responding to chemotherapy (13).

1.2 Lung Cancer Progression

Different lung cancer subtypes arise at different anatomical locations. Though multiple anatomical models of the lung have been described, the most common is the Weibel version (14). This model numbers successive generations of the airways starting at the trachea (generation 0) down to the alveolar sacs (generation 23). Full details of the structural anatomy of each generation are shown in Appendix A Table A.2.

Between generation 11 and 12 is the transition from the central airways to the peripheral airways. Distinction between these two regions is important. Cartilage is absent beyond generation 11, tissue layers become thin, and peripheral airways have a much smaller diameter than central ones. This restricts access to the epithelium, which results in significantly reduced detection strategies. As well, lung cancers that arise in the periphery are often of a different subtype than central ones. As such, this thesis discusses the central and peripheral airways separately.

1.2.1 Lung Cancer in Central Airways

The most common lung cancer subtype in the central airways is SCC, which is a multistage disease (15). The sequence of events in the development of SCC is 1) carcinogens cause injury to the DNA of cells, 2) the injury is too great for the cell to repair, 3) the normal cell
cycle mechanisms malfunction which leads to an accumulation of faulty cells causing a morphological change.

Accepted carcinogenesis models of SCC suggest that the most significant carcinogen in the central airways is cigarette smoke (16-19). Tobacco smoke contains polycyclic aromatic hydrocarbons (PAHs), which have significant impacts on the formation of lung cancers. These chemicals generate the production of reactive oxygen species (ROS), which are unstable forms of oxygen that are far more reactive (16). The main targets for ROS damage are DNA, lipids and proteins. Over the course of frequent exposure to ROS normal cell defense mechanisms used to repair damaged DNA (and other biomolecules) are overwhelmed and eventually fail resulting in genetically damaged cells giving rise to damaged progeny (16, 20). Tobacco carcinogens also bond directly to DNA causing various genetic abnormalities directly. The second most common carcinogen in the central airways is radiation from inhaled Radon gas (16). Alpha particles (He) can cause direct DNA damage or again create ROS. Finally, there is a subset of male patients who are never smokers presenting with SCC (female never smokers are more likely to be diagnosed with Adenocarcinoma) (8). In these cases, an environmental carcinogen such as arsenic may be the initiating factor of tumor progression. Inherited gene mutations, epigenetic malfunctions, infections and diet have all been cited as lung cancer risk factors (8). In general, the number and severity of genetic abnormalities correlate with the histopathological grading. Low grade lesions have far fewer genetic abnormalities than high grade ones.

Once a cell has evaded the normal cell cycle processes, it will replicate and produce progeny containing the genetic damage. Many histological studies have shown that there is an association between early cell proliferation in the mucosal layer of large airways and subsequent diagnosis of SCC. Tumors will progress through hyperplasia, metaplasia, dysplasia and
carcinoma in situ (CIS) before becoming an invasive disease. The natural timeline and progression of these pre-invasive tumors have been studied extensively and it is known that they often regress back into a normal state (15). However, those lesions that do progress beyond the dysplastic state have a higher probability of becoming SCC. Full pathology details can be seen in Appendix A Table A.3.

Proliferation of cells close to the basement membrane causes a disorderly arrangement of normal cells above towards the luminal surface (21-23). As this proliferation increases, and cells stop maturing, the disorder towards the luminal surface increases. In the next phase, the cells become invasive and break through the basement membrane (21-23). Although many hyperplasias, metaplasias and dysplasias will regress, there is strong evidence that high grade dysplasia (HGD) is a risk factor for the development of Invasive Carcinoma (IC) (15). Detection of these premalignant (HGD) lesions is important and determines which patients require a close follow up.

However, detecting these pre-invasive lesions is very difficult because they present very few changes macroscopically. The ideal detection technology in the central airways would be able to probe below the luminal surface and characterize pre-invasive and invasive pathologies.

1.2.2 Lung Cancer in Peripheral Airways

The most common lung cancer in the peripheral airways is Adenocarcinoma (ADC). Although the accepted sequence of events leading to malignancy is the same as SCC, there is debate over the initial cause of increased cell proliferation. Lung scarring was once thought to be a significant precursor to peripheral adenocarcinomas – as the tissue repair pathway calls for cell proliferation as well. However, tissue scarring has now been shown to be a secondary event
in the malignancy pathway (24-26). Others have suggested that carcinogens from cigarette smoke remain highly relevant because dosages from nitrate levels in tobacco smoke have increased (8, 18). Nitrate is able to form nitrosamines, an ROS producing chemical associated with the development of adenomas and ADC. As well, changes in filters and inhalation patterns of cigarette smoke cause deposition of particles deep into the peripheral airways (8). Finally, a subset of female never smokes present with ADC (8). Striking risk differences among global populations have been described, showing much higher occurrence and risk for females living in East Asian, particularly in China. Second hand smoke has been suggested to be the most significant risk factor, along with underlying genetic susceptibility and environmental outdoor and household pollution (8).

Pre-invasive lesion description in the peripheral airways is relatively new, but a multistep progression analogous to SCC has been proposed. Atypical Adenomatous Hyperplasia (AAH), Adenocarcinoma in situ (AIS), minimally invasive Adenocarcinoma (MIA) and Invasive Adenocarcinoma are now commonly accepted as the progression pathway in the lung periphery (27). Although the natural timeline of these lesions have not yet been fully elucidated, there are a number of studies that support the association of AAH and AIS with ADC (27).

Due to the difference in tissue layers and structure between the central and peripheral airways (Appendix A Table A.2), there is also a difference in lesion progression. AAH is a small localized proliferation of atypical cells (type II pneumocytes or Clara cells) that line the alveolar walls. As proliferation increases, and cells become more atypical, a lesion will spread along the surface of the alveolar wall (lepidic pattern) (27). Further proliferation results in the invasion of the alveolar cell wall but if this invasion remains limited and tumor size is <3 mm, the five year survival rates are extraordinarily good – studies showing 100% five year survival
rates have been reported (27). Thus, the detection of these early or precursor lesions is incredibly important.

Similar to SCC, symptoms often appear late, beyond the AIS or MIA state. However, different than SCC abnormal tissue in the periphery often arises in the alveolar spaces, not on the surface of the small airways – compounding the challenge of detecting lesions early and highlighting the need for technology that probes beneath the tissue surface. Further, the structure of the peripheral airways themselves causes a significant challenge for early detection as the small diameter physically reduces access to the epithelial surface. Detection strategies in the periphery, discussed in detail in Section 1.3, have high requirements. Technology used must be able to work in small diameters, determine differences between benign (non-cancerous) and malignant conditions, and recognize the malignancy pathway in the early stages.

1.3 Current Clinical Diagnostic Methods

In both the central and peripheral airways, the earlier a cancer is discovered the better the prognosis and a more favorable outcome there is for the patient (28-30). However, a definitive diagnosis of a lung lesion is often required to proceed in the treatment pipeline. Therefore, if a suspicious lung lesion is found by a physician, the gold standard of diagnosis is a biopsy followed by histopathology. If necessary, lymph nodes will also be biopsied to determine staging. Thus, the diagnosis of lung cancer requires 1) the detection of suspicious lung lesions 2) the localization of lesions to take a biopsy, and 3) the retrieval of good tissue sample from that location. These three steps are discussed in detail below.
1.3.1 Detection Methods

1.3.1.1 Whole Lungs

Historically, the detection of lung cancer in the central and peripheral airways was through chest radiographs and analysis of sputum samples. However these techniques are unable to detect early or premalignant lesions. As discussed in Section 1.2, SCC and ADC start as non-invasive precursors. The small size of precursor lesions allows for evasion of detection by chest x-rays which have been shown to have low sensitivity (9, 31). Nodules that appear on chest x-rays have already progressed into invasive disease. Further, these precursors cause relatively little structural damage and are contained within the luminal layer (CIS) of tissue or contained along the surface of an alveolar wall (AIS), meaning tumor bi-products and cell sloughing are either contained, or relatively small in volume, which evades sputum testing (32). The sensitivity of sputum cytology for lung cancer screening is between 20-30% (32), thus, these techniques have been declining in use for more favorable ones.

A major clinical study in the last decade compared the use of chest radiographs with low-dose CT (LDCT) as a cancer screening method. The National Lung Cancer Screening Trial (NLST) (9) found that mortality rates of lung cancer could be significantly reduced by screening with low-dose CT when compared to a chest radiograph because LDCT can detect early stage and pre-invasive cancers (9, 33). As such, the use of low-dose CT for screening has been implemented into clinical practice. CT scans take multiple X-ray image slices horizontally across the body, which are then stacked to produce a 3D image that a physician can use to view the entirety of the lung. The thickness of slices can be altered depending on the clinical need. Thin slices are able to produce a more complete 3D image (as each horizontal slice would be almost adjacent to each other) but there is a trade-off in the amount computational power needed.
to store many images and the specialist requirement of having a physician review all the images. In general CT has a resolution capable of imaging 0.5 mm lesions (34).

Yet, CT scans cannot definitively characterize a lung nodule as benign or malignant. As well, the pitfall of extremely high sensitive LDCT for lung cancer screening is an equally high false positive rate (9, 35). Further, once a lesion is found using CT, that lesion still requires a biopsy for a definitive diagnosis. Since LDCT is imaging the lung globally and externally, localization of the suspicious tissue is still required in vivo. This process is notoriously difficult in the peripheral airways.

1.3.2 Localization and Biopsy Retrieval Methods

1.3.2.1 Central Airways

Bronchoscopy

One of the major breakthroughs in lung imaging in the last century was bronchoscopy, which allows clinicians to visualize the surface of the airways. Bronchoscopy has been used for the last two decades for localizing early stage cancers of the central airways, during which time there have been many technological improvements to the bronchoscope itself, and the development of helpful adjunct devices (34, 36-40). Specifically, auto-fluorescent bronchoscopy (AFB) was developed and increasingly used in adjunct to the standard white light bronchoscopy (WLB) to improve the localization of early stage cancers and HGD of the main airways (36, 37, 41, 42). A recent meta-analyses of 21 studies performed by Sun et al (43) showed that combined WLB+AFB exam greatly improved the sensitivity for localizing HGD or CIS (premalignant lesions), but with decreased specificity as compared to WLB exam alone. The pool relative sensitivity of WLB+AFB versus WLB was 2.04, while the pool relative specificity
of WLB+AFB versus WLB was 0.65. This highlights the need of developing new adjunct technologies for improving the specificity. Premalignant lesions are a risk factor for the development of IC somewhere in the lung (15), so identifying lesions in the central airways with HGD provides valuable information about patients requiring close follow-up. There was however considerable variation in the sensitivities (41-100%) obtained across all 21 studies that used a combined WLB+AFB for indentifying premalignant lesions. The largest variability in sensitivities occurred at high specificities which probably reflect differences in biopsy protocols, and operator experience. Those data suggested that in order to increase the probability of achieving a high sensitivity a liberal biopsy protocol is required otherwise the risk of missing a premalignant lesion escalates significantly. The drawback with such an approach is that it will result in many biopsies which are negative for premalignant lesions or worse, which may lead to longer procedural times, greater healthcare costs and a higher incidence of procedure complications. These factors hinder the practical adoption of this technology for widespread clinical uses. The same meta-analysis also showed that the specificity obtained across all 21 studies varied from 18-86%. The general trends are that for studies achieved high sensitivities with WLB+AFB, their specificities were relatively low.

**Other Techniques Used Adjunct to Bronchoscopy**

Various other optical methods have been shown to have value as well. Optical coherence tomography (OCT) was used as an adjunct technique to WLB+AFB to predict the pathology of lesions located with AFB. OCT, a method utilizing light wave scattering and reflection similar to an ultrasound, is able to show physical structures of the lung (34, 44, 45). As well, narrow band
imaging (NBI), an endoscopic method which illuminates mucosa at a wavelength of optical absorption for hemoglobin, is able to show physical features of blood vessels (34, 46, 47).

A number of other technologies have been explored to improve the specificity of an AFB + WLB examination. The ratio of red to green (R/G) fluorescence of a lesion is the easiest to implement requiring minimal additional equipment, but so far the results have been mixed (48-50). Another technology that has been explored to improve the specificity in identifying HGD during a bronchoscopy is reflectance spectroscopy. However, initial in vivo tests were conducted on a relatively small number of lesions which had either a benign/normal or IC pathology; very few, if any HGD were included (49).

1.3.2.2 Peripheral Airways

Two methods are used for the localization and collection of biopsy samples from the peripheral airways. The first, the CT-guided Transthoracic lung biopsy (CT-TTLB) procedure, uses live CT imaging to direct an external hollow-core needle into the lesion location. The yield of the procedure is relatively high, where 70-90% of the time diagnosable biopsies are obtained (51) and lesions as small as 0.5 cm can be biopsied with this method (52). However, this procedure carries the risk of serious adverse events like pulmonary bleeding and pneumothorax. Some studies have reported post-procedure pneumothorax events to be as high as 17% (51). These adverse events also increase in severity with pulmonary emphysema the length of the needle used and the number of repeat CT-TTLB procedures performed (51).

The second method used is endoscopic transbronchial lung biopsy, taken from the luminal surface of the peripheral epithelium. Standard WLB+AFB cannot be used in the
peripheral airways for localization because the airways have a diameter that is too small and beyond the reach of conventional bronchoscopes. Here, physicians use Radial Endo-bronchial Ultrasound (R-EBUS), a procedure that extends the length of the instrument channel in a bronchoscope and localizes which lobe and division nodules are in (34). A small rotating transducer attached to the distal tip of a probe is sent down the instrument channel of a bronchoscope. Similar to other ultrasound technologies, the small transducer sends out high frequency sound waves which are then echoed by physical structures of the body and collected by the same transducer. Differences in density of tissue show as distortions on the ultrasound image. Tumors or lesions that are located either at or behind the luminal surface of the lung are able to be detected because of the high penetration depth of R-EBUS (4-5 cm range) and spatial resolution between 200 µm and 1mm (53, 54).

When compared to CT-TTLB, the R-EBUS procedure has significantly less risk of adverse events, but the diagnostic yield for peripheral lung lesions is lower, around 60% (55, 56). One possible reason for the lower yield is the necessary removal of the ultrasound probe prior to biopsy collection which can create displacement of the guide sheath. Further, the procedure cannot characterize benign and malignant conditions (53). Large blood vessels cannot be distinguished from nodules, and non-solid nodules like ground glass opacity are very difficult to visualize with R-EBUS. As well, the forceps used with R-EBUS are end-grabbing, where the R-EBUS probe is side-viewing so the collection of tissue does not precisely match the location indicated by the navigation with R-EBUS.
1.4 Clinical Gaps

1.4.1. Central Airways

The ease of access and visualization of the central lung epithelium along with the inclusion of WLB + AFB as a routine clinical procedure has revolutionized the localization and subsequent biopsy of pre-invasive lesions in the central airways. Current equipment and technologies have a high sensitivity but the specificity of WLB+AFB varies widely (43). In general, the specificity is low when high sensitivity is achieved. This low specificity is due in part to the training of the user and the user’s familiarity with the color scheme of the current equipment as device manufacturers have not standardized technology platforms (57). Further, false-positive changes in tissue that are unrelated to cancer development, like inflammation and trauma, can contribute to the low specificity of WLB + AFB. There is an opportunity to improve the specificity of WLB+AFB for the localization of premalignant/malignant lesions in the central airways by providing an objective quantitative measure of lesion characteristics to reduce variation among WLB + AFB users (57).

1.4.2 Peripheral Airways

The increased use of LDCT for lung cancer screening has resulted in frequent detection of small lung nodules, many of which are located in the peripheral airways. Coupled with a growing trend of increasing ADC and decreasing SCC, peripheral airway nodules are becoming more relevant. These nodules need to be localized and subsequently sampled for biopsy to confirm or deny a lung cancer diagnosis. The least invasive method to localize and biopsy these nodules are through Radial Endo-Bronchial Ultrasound (R-EBUS), where the diagnostic yield is relatively low at about 60%. Here, the ability to characterize the nature of tissue in real time
before taking a biopsy would allow for more precise navigation to premalignant and malignant tissue.

1.5 Raman Spectroscopy

The phenomenon of Raman Scattering was initially described in 1928 by Sir C.V Raman, an Indian physicist who subsequently won the Nobel Prize in 1930 for his work (58). Yet, recent advancements in optic and photonic technologies have enabled the manipulation of light to obtain clinically relevant information in a rapid and minimally invasive way. The availability of optical components such as fibre optics, sensitive cameras and diode lasers has allowed researchers to bridge the gap between experimental setup and clinical implementation. The clinical applications of optics, techniques based on fluorescence and absorption in particular, have continuously claimed new frontiers in the last two decades (37, 59-61). Researchers are now able to elucidate different disease states, make diagnoses and optimize treatment for a patient based on understanding the fundamental interactions between photons and tissue. At the same time, maturation in technologies such as lasers, gratings, and CCD cameras have largely overcome limitations of conventional Raman spectroscopy (RS) and made its biomedical utility a possibility. Raman spectroscopy, probes molecular vibrations and gives very specific fingerprint-like spectral features (62) making it useful for biochemical bond identification, and thus classification of biological tissues.

The theory of Raman interactions and other various light-tissue interactions are well understood. When a photon interacts with tissue, there are many possible outcomes. Reflectance, absorbance and fluorescence are all possible – though limited biochemical information is retrieved from these events and they are often utilized to obtain morphological or
physiological characteristics about tissue. Yet, if a photon is scattered by a molecule, there is the possibility of obtaining specific biochemical information about the bonds which are present in the sample.

If these scattering events do not change the frequency and energy (wavelength) of the photon, they are known as elastic scattering, or Rayleigh scattering. Scattering events that do alter the frequency and energy of the photon are known as inelastic scattering, or Raman scattering, occurs when an incident photon is scattered by a molecule and emerges at a wavelength different from that of the incident photon. This wavelength difference is related to the energy required to excite a corresponding molecular bond into a different vibrational energy state as shown in Figure 1.1. When the photon transfers energy to the chemical bond in Raman scattering, a Stokes shift in wavelength occurs, while an Anti-Stokes shift in wavelength occurs when the photon gains energy from the bond (63-67). Stokes scattering is a more likely occurrence because the vast majority of molecules are in the ground vibrational state (68).

When light is scattered by a molecule most photons are elastically scattered (Rayleigh Scattered ), and only approximately one in a million photons is inelastically scattered and emerges with a change in energy (69). Recent work has demonstrated that, though relatively small in numbers, these Raman scattered photons can be quantified with reasonable statistical accuracy, at incident light intensities that do not damage even the most delicate of samples, and with relatively short exposure times. In simple pure molecular samples the measured intensity of the Raman scattered photons from one vibrational mode is proportional to the abundance of those molecules in the sample (69). For biological samples, where many different types of molecules can contribute to the measured Raman signal, the relationship between the intensity of a Raman peak and molecular abundance is complex since many types of molecules can share the
same chemical bonds, and different chemical bonds can share the same energy. Nevertheless the variation in the wavelength position and intensity of all the Raman peaks in a Raman spectrum is highly specific to a unique set of chemical bonds.

**Figure 1.1:** Schematic showing the difference between Rayleigh, Stokes and Anti-Stokes Raman Scattering. When a difference in energy occurs between the initial vibrational energy state and the resulting vibrational state, a shift in the frequency of the scattered photon will occur. The incident photon has wavelength $\lambda_e$ and the inelastically scattered photon $\lambda_s$. The difference between the inverse of these two wavelengths is known as the Raman Shift, a value that is directly related to energy. Reproduced with permission from (1).
Raman scattering can occur for all photon wavelengths that make up the incident light, provided there is a vibrational mode in the sample with transition energy less than the energy of the incident photons with the longest wavelengths. In practice, the most information about the sample can be obtained from Raman scattering when using incident photons with a very narrow peak shaped band of wavelengths. In this case the Raman scattered light from one vibrational mode is also emitted in the shape of a peak, called a Raman peak. Thus knowing the centre wavelength of the incident light and measuring the centre wavelength of the Raman peak, the vibrational energy of a molecular bond can be determined. The narrower the band of incident wavelengths (excitation light that is monochromatic is desirable), the better it is for separating Raman peaks caused by different vibrational modes that are separated by a small energy difference. Raman Spectroscopy utilizes Raman scattering by probing a wide range of molecular vibrational energy modes of a sample either sequentially or simultaneously which is more typical of modern systems. The output of this probing is a characteristic series of Raman peaks called a spectral fingerprint, containing information about the chemical structures and quantities of specific molecules in the sample. In other words, a Raman spectrum is a function plot of scattered intensity (photon counts) versus the frequency difference of incident and scattered photons, known as wavenumbers.

Clinical Raman spectroscopy measurements can be accomplished through macro samplings (≥ mm$^3$) of tissue or bodily fluids where the signal mainly comes from proteins, lipids, and nucleic acids, within the excited tissue volume. Changes in tissue pathological states, which are usually preceded by and correlated with biochemical changes, can be inferred through spectral analysis, demonstrating Raman Spectroscopy to be a valuable clinical tool (70-72). On the other hand, if initial neoplastic growth begins within epithelial layers of the tissue, macro
sampling may introduce unwanted Raman and endogenous autofluorescence (AF) signals generated from deeper tissue layers which can overwhelm the clinically relevant Raman signal. To overcome this problem, hardware components are often selected to minimize the contribution of AF. Since AF is usually most severe with shorter excitation wavelengths, using a different excitation wavelength or by detecting the Raman in the high frequency region (1800-3200 cm\(^{-1}\)), the endogenous AF can be reduced enough to extract the Raman peaks. Although, Raman peaks in the fingerprint spectral region (\(\approx 500-1800\) cm\(^{-1}\)) tend to have more biochemically associated signatures, so there is a trade-off. As AF is reduced so is the amount of Raman information available. These two wavenumber regions can be measured simultaneously with the same equipment or probed separately at a higher spectral resolution through hardware modifications such as a change of excitation wavelengths or the spectrometer configuration.

Raman Spectroscopy has previously been studied in clinical settings, and has shown use in classifying skin lesions (62, 73), use in the cervix (74), colon (75) and oral cavities (4). The only adaptation of a clinical Raman system designed for the lung was by our group (76). The study showed Raman spectroscopy measurements were able to classify benign from malignant lesions in the central airways with high sensitivity and specificity. However, this study was a small pilot study and consisted of only the central airways. There is a need to validate the findings using a large clinical study and assess the feasibility of using RS in the peripheral airways.
1.6 Thesis Aims and Outline

It is hypothesized that real time rapid Raman Spectroscopy can be introduced into clinical procedures and is capable of detecting the biochemical differences among normal and various abnormal lung tissue, which can then be used to classify tissue in real time and detect premalignant and malignant lesions in the lung. This thesis is structured to test this hypothesis by addressing the following aims.

This thesis aims to 1) validate the performance of RS for in vivo premalignant and malignant lesion detection in the central airways through a clinical study and statistical analysis, 2) design a novel miniature Raman probe for use in the peripheral airways and 3) determine the feasibility of using RS for in vivo lesion classification in the peripheral airways.

To address aim 1) I present, in Chapter 2, the hardware design of our in house clinical lung Raman system. Though I was not the creator of this system (credit goes to Dr. Michael Short and Dr. Haishan Zeng of my lab), I explore ways to improve on the signal to noise ratio of the system, particularly by hardware improvements. By comparing two ex vivo experimental data sets collected on the old and new RS systems, I can demonstrate the improvement of hardware results in higher quality spectra. Further, by completing a full statistical analysis on an in vivo data set of oral tissue I can demonstrate the new system’s ability to differentiate tissue in humans.

To further address aim 1), in Chapter 3 I present and discuss a large scale clinical trial in which we explored the performance of using RS on improving the specificity of localizing premalignant and malignant lesions in the central airways. I collected in vivo Raman spectra from 80 patients at the BC Cancer Agency (Vancouver Centre). I measured a total of 280 sites; 72 were either a premalignant or a malignant lesion (CIS or IC) and 208 were either a benign
lesion or a normal tissue site. The spectra were analysed using different multivariate statistical methods, achieving a high sensitivity and specificity.

To address aim 2), I present in Chapter 4, the design and developed novel fabrication procedure and initial testing of a miniature Raman probe for the peripheral airways. I designed a novel fabrication procedure that facilitated the creation of a probe flexible enough to reach the peripheral airways. Through bench top measurements on Tylenol and in vivo measurements on the skin, I demonstrate that the miniature Raman probe has good collection efficiency and functions optically similar to the central airway probe.

Finally to address aim 3), I present in Chapter 5 and 6, the initial clinical tests of the miniature peripheral Raman probe, where I show the first, to my knowledge, in vivo peripheral Raman measurements. I also explore the feasibility of using the designed probe to classify benign peripheral lesions outside the scope of lung cancer.
Chapter 2: Optimization of Hardware and Data Analysis Methods

2.1 Chapter Introduction

Developing new clinical technologies based on Raman Spectroscopy requires a high signal to noise ratio (SNR). The inherently weak Raman signal (due to the small number of Raman scattered photons) is often difficult to collect and detect. Further, other confounding modalities such as fluorescence can overwhelm the weak Raman signal and mask it. These background signals often need to be mathematically removed prior to developing a classification algorithm – which too needs to be optimized in order to get the most relevant information possible from the tissue. Therefore, this Chapter addresses the development pipeline of a Raman-based clinical system. First, hardware components and instrumentation are analyzed for the best SNR. Second, mathematical techniques and classification algorithms are developed from small ex vivo and in vivo data sets. The goal of this Chapter is to determine the best components for a clinical system which can collect enough data for a large clinical study (Chapter 3), and once that data is collected, to determine the best way to elucidate the information it holds.

2.2 Instrumentation

In 2008 the first Raman based system for use in lung endoscopy procedures was developed by our lab (77). In 2011 this system was used to conduct a pilot study, the results of which showed Raman was an effective way to differentiate between pre-malignant and malignant lesions (76). This warranted a large clinical trial as follow up to validate the findings, expand the classification algorithm, and start the process of regulatory approval. However, prior to investing resources into such a large undertaking, it was prudent to make sure the data
collected in a large clinical trial would be of the highest quality. Changes in the system were made with updates that included the latest optical technologies. These improvements were then tested and analyzed against the old system (the first generation system) (2).

2.2.1 Old versus New System Design

The old Raman system (76) had a spectrograph coupled to a charge coupled device (CCD) (model: Spec-10:400BR/LN, Princeton Instruments, Trenton, NJ) that was cooled through the use of liquid nitrogen to about -95°C, was optimized for near infrared wavelengths, and was deep-depleted and back-illuminated. The spectrograph was an f/2.2 (model: HS-f/2.2-NIR, Kaiser Optical Systems, Ann Arbor, MI) system with a holographic grating (HSG-785-HF) fixed in the high frequency wavenumber range from 1570-3400 cm\(^{-1}\). Further, the spectrograph had a numerical aperture (NA) of 0.22 which was well matched for the NA of the input fibre which was 0.22.

Different from this old Raman system, the newly updated system (2) included a spectrograph coupled to an improved charge coupled device (CCD) (model: PIXIS 400BR, Princeton Instruments) which was cooled thermoelectrically to -70°C and had a better etaloning reduction architecture (78). The spectrograph was a f/2 (model: LS-785, Princeton Instruments, Trenton, NJ) system with a holographic reflection type grating that was able to be manually tune a \(\approx 2000\) cm\(^{-1}\) wide spectral window over a wavenumber range from 0-3400 cm\(^{-1}\). This spectrograph was also equipped with anti-reflection (AR-coated) lenses for optimized near infrared (NIR) transmission. Further, the NA of the system (0.24) was also well matched to the NA of the input fibre (0.22).
The tissue excitation was induced similarly in both systems with a wavelength stabilized 785 nm diode laser (model: BRM-785, B &W Tek, Newark, DE). A probe was designed to work specifically on the lung mucosal surface. It consisted of a single, ultra-low OH (impurity), 200 µm diameter fibre (which was gold-jacketed to prevent fibre cross-talk) for excitation surrounded by 27 collection fibres which were also ultra-low OH, but 100 µm in diameter to create a circular fiber bundle of 1.8 mm in diameter (same as the old system). In the new system, 3 collection fibers were sacrificed for the Visible Light Channel as some existing equipment cameras cannot detect the laser spot on the tissue surface; therefore the visible light channel was needed for the alignment of the probe to the lesion, as shown in Figure 2.1. The excitation power was 150 mW at the mucosal surface with an illumination spot size 3.5mm in diameter (120 mW in the old system). The calculated laser power density is within the American National Standards Institute (ANSI) maximum permissible exposure limit for the skin which is 1.6 W/cm² (79). The entire system was controlled using a PC through in-house created software (80). The software included an algorithm to remove the background tissue autofluorescence in real time (80).
Figure 2.1: Schematic diagram of the new *in vivo* Raman system (not to scale), with magnified views of the fiber tips. The old system was similarly organized except it did not include the Visible Light Channel.

### 2.2.2 Standard Comparison

The performance of the two systems was compared by measuring a N-acetyl-para-aminophenol (Tylenol) standard with a fixed probe to sample geometry. The same fiber probe and in house software were used for both systems. 100 spectra were obtained of both the sample and the dark background. The power difference in the systems was then corrected for by multiplying the power ratio into the old system (which had a reduced power). The mean of each collection is shown in Figure 2.2. The signal to noise ratio was then determined for each system and compared, as shown in Figure 2.3. This was done first by removing the background of the
signal spectra by the subtraction of an averaged (n=100) spectra where the laser was off. The standard deviation of the spectra was calculated and then this was divided into the average intensity to generate the signal to noise ratio.

Figure 2.2: The average of 100 Raman Spectra of Tylenol using both the new and old system.

It should be noted that the fluorescence has not been removed and the spectra have been corrected for power differences.
Figure 2.3: The signal to noise ratio of both the old system and the new system (when compared using the Tylenol standard) corrected for power differences.

The signal to noise ratio compares the average number of (Raman) emitted photons (signal) to the standard deviation in that number. This standard deviation is comprised of the statistical fluctuations in sample emission and system generated electrical signals (noise). Therefore, the signal to noise ratio was compared between the two systems, to determine the system with the higher signal to noise ratio. This method showed that there was a higher signal to noise ratio across the measurement range for the new system, consistent with our observations.
that the fluctuations in system generated noise are reduced and the sensitivity of the system is increased.

The improvement in the signal to noise ratio can be attributed to the improved spectrograph and CCD on the new system. The Princeton spectrograph is equipped with AR coated lenses and improved optical components like a gold coated grating for high optical transmission. Further, the PIXIS CCD is run at a slightly higher temperature (\(-70^\circ\text{C}\) compared to the old system that runs at \(-95^\circ\text{C}\)) which improves the quantum efficiency of the camera. The slightly higher temperature increases the energy of the CCD electrons which results in lower energy needed to create an electronic signal. Therefore, the inherently weak Raman signals can be detected with higher sensitivity in the PIXIS CCD.

2.3 Spectral Pre-processing Techniques and Multivariate Statistical Methods

2.3.1 Spectral Pre-processing Techniques

Raman spectra are often collected with other modalities, such as NIR tissue autofluorescence, so removal or reduction of these unwanted signals prior to classification is necessary as they can be confounding. Below is a brief overview of the pre-processing methods used throughout this thesis.

Fluorescence Background Removal

The Raman signal is usually superimposed on a large broad fluorescence background so the most common removal technique is a polynomial fitting. This technique is convenient and simple to use. The background is modeled by a polynomial function whose order is chosen for
the best fit (usually a 5th order polynomial), and then subtracted from the spectra leaving the
Raman signal alone (81). However, variability often arises from this method as orders and fitting
ranges need to be chosen by the researcher and is often arbitrary. To combat this, in 2007 Zhao et
al. from our lab constructed an automated autofluorescence background subtraction algorithm
based on a modified curve fitting (the Vancouver Algorithm), the details of which can be seen in
ref. (80). The Vancouver Algorithm (80) is an iterative multi-polynomial fitting method that
combines polynomial curve fitting, peak removal and noise reduction (smoothing). An example
of fluorescence background removal using the Vancouver Algorithm (80) on *in vivo* biological
data (palm skin) can be seen in Figure 2.4. Two significant benefits of the Vancouver Algorithm
are first, that it is capable of functioning in real time and it is therefore desirable for clinical
applications of Raman Spectroscopy. Second, the Vancouver Algorithm (80) is capable of
providing relatively pure Raman spectra which can then be analyzed in a multitude of ways.
Throughout this thesis, the Vancouver Algorithm is the primary method used for the
fluorescence background removal of biological spectra.

Another method used for fluorescence removal is the calculation of the second order
derivative (76). Our lab has previously used the Savitzky-Golay polynomial to calculate the
smoothed second order derivatives of spectra, details of which can be seen in ref. (82), and a
detailed example of the method can be seen in Appendix B. The second derivative is the rate of
change in the slope of the spectra (the curvature of the spectra). An important property of
derivatives is that peak width can affect the amplitude of a derivative so relatively small but
sharp Raman peaks on a broad background can be extracted with some precision. This precision
allows second order derivatives to accentuate small structural differences in the spectra even
when the fluorescence background is very high. However, the main limitation to using second
derivatives is that the pure Raman spectra shape is complicated to obtain. Therefore, comparing the spectral shape to published literature is difficult.

**Figure 2.4:** The fluorescence background removal using the Vancouver Algorithm (5\(^{th}\) order polynomial) on an example Raman spectrum from *in vivo* palm skin.

**Normalization**

After fluorescence background correction, spectra must be normalized. This will account for inter- and intra spectral variability such as probe distance to the sample and power fluctuations. Though techniques such as peak normalization and standard normal variate transformation are used (83, 84), one of the most common normalization techniques used is
normalization to the area under the curve (AUC) (62, 76). In this process, the area under each curve is summed and then each variable in the smoothed spectrum is divided by this sum.

**Spectral Peak Analysis**

Once the spectra have been pre-processed, the pure Raman spectra can be compared to published literature. Often, the location of a peak or specific waveband can be assigned to a vibrational mode of a bond or molecule. Common literature databases like the one published in ref. (85), allow for comparison of the retrieved signals versus biological standards, giving insight into the bio-molecule or biochemical that is related to the Raman peak.

**2.3.2 Multivariate Statistical Methods**

Raman spectra contain large amounts of information (multiple wavenumbers) so multivariate statistical techniques are often used (5, 62, 76, 86, 87), as a single peak or wavenumber often cannot provide the necessary discrimination between tissue types. There are many multivariate techniques that have been shown to be successful in classifying Raman spectra. General techniques are introduced and discussed briefly below. Each subsequent Chapter in this thesis also specifies the statistical methods, procedures, software and programs used that are relevant to the Chapter. A flow chart of techniques used throughout this thesis can be seen in Figure 2.5, and detailed information about the multivariate techniques can be seen in Appendix C.
Waveband Selection

A single Raman spectrum contains a large amount of data. Often over 1000 wavenumbers are collected, not all of which contain important information that can be used for tissue discrimination. Selecting the wavebands or the regions of the spectra that contain significant information have been shown to improve prediction and classification models (88-92). Three techniques will be discussed here. The first, stepwise multiple regression (STEP), is a systematic method for constructing a multilinear regression model (93, 94). The STEP algorithm computes the P values (of the F-statistic) of each wavenumber between two groups, and then the algorithm adds/removes wavenumbers from the model based on these P values. The algorithm will start by adding the most significant wavenumber (smallest P value of the F-statistic) into the model, and subsequently refit the model. It will then continue adding wavenumbers - the most significant one below a pre-defined P value cutoff. For example, if the cutoff value is set at 0.01, the smallest P value below 0.01 will be entered into the model. Subsequent recalculations of the model occur until there is an optimum model determined. Often, wavenumbers will also be removed from subsequent models based on the refitting (i.e. the wavenumbers which have a P value that is higher than the pre-defined cutoff will be removed). Therefore, the algorithm will stop when the P value (of the F-statistic) of all the wavenumbers in the model meet the entrance and exit value requirements. The simplicity of STEP makes it an attractive choice for waveband selection, especially on data sets with a very large amount of wavenumbers. Although it is limited by the initial statistics of the data set as the most significant wavenumber is the first to be included in the model. Further, it is a localized method (not a globalized method) so the optimal model found may not be the best model for the data. Yet, this localization also means that
computational power is reduced because the method does not search through all possible subsets of the data.

The second method, Least Absolute Shrinkage and Selector Operator (LASSO) is a penalized regression technique (95). The LASSO, is an embedded method that develops a linear model but applies an L1 penalty to many of the model coefficients. This L1 enforces a maximum size on the sum of the absolute value of the magnitude of the coefficients, which causes many of the coefficients to shrink to zero. The non-zero coefficients remain in the model. This method essentially chooses a reduced number of coefficients to be included in the model, effectively reducing the complexity of it. Though similar to the STEP method, LASSO is statistically driven and thus wavenumbers will be included (or not included) in the model based on significant P values.

The third method, the Genetic Algorithm (GA), is a method that finds the optimal model of the best subset of data (96, 97). The GA is an iterative method that calculates the P value at each wavenumber and any that are below a set threshold will be included in the first generation model. These initial wavenumbers are then used to create all possible predictive models of the data. Subsequent iterations will take wavenumbers in the current generation at random to create the next generation model – where again all possible combinations are used to produce a population of predictive models. This process is repeated until the algorithm “evolves” into the optimal model. Since all combinations of wavenumbers are created and analyzed, the GA is in some sense a globalized method. However, the method is not repeatable since random wavenumbers are chosen to create the next generation. Often this will be addressed by running GA multiple times and then taking the average of all the results.
Dimension Reduction

Following waveband selection, dimension reduction techniques are often used for feature extraction. These techniques will transform the data into a space with fewer dimensions, allowing for the extraction of important prediction features. Two techniques will be discussed here. The first, Principal Component Analysis (PCA), is an orthogonal transform that maps the original data into a lower dimension space where the variance among predictive features is maximized (98). Here, linear combinations that explain most of the variance in the data are extracted. These are known as Principal Components (PCs) and are related to the original spectra by a weighted eigenvector known as the PC score. Each PC score has a corresponding eigenvalue. The larger the eigenvalue of a PC score, the more variance that is contained within that particular Component, and thus, top scores will often be used in classification models to separate two groups. PCA is used frequently in Raman spectral analysis as the computational power needed is minimal and it is a straight forward process. However, problems may arise in data sets that have a large amount of noise – as the top principal components accounting for the majority of the variance may just be detecting this noise. Further, PCA maximizes the variance among the predictive features only, and no importance is given to the classification of those features.

The second method, Partial Least Squares (PLS) is similar to PCA, with slight variations in that it is an orthogonal transform that maximizes the variance between groupings of the predictive features. In other words, it is a dimension reduction method that also maximizes the separation of classes among the data set. In practice this small difference can play a large role in the performance of a classification algorithm and our lab has shown that PLS often outperforms PCA on analysis and classification of biological data (5, 91).
Classification Models and Validation

Effective classification models for discriminating different Raman Spectra aim to minimize the variance between members of the same group, while maximizing the variance between members of different groups. Linear Discriminant Analysis (LDA) will determine a function line that can be used to predict the classification of new cases (99). Similarly, General Discriminant Analysis (GDA) will also determine a function used to predict new cases, but is the generalized form of LDA. The performance of these models are validated in numerous ways, the most common being the Leave One Out Cross Validation (LOO-CV). Here, one spectrum is removed from the entire data set and an optimized classification model is developed with the remaining data. The withheld spectrum is then tested against the developed model to calculate a probability of being in one of the classification groupings (termed posterior probability) (100). This process is repeated until each spectrum is the one left out and classified with a posterior probability (101).

ROC Curves

A random classification of a spectrum has the probability of 0.5 for being classified in one of two groupings. From the probability plot, an ROC curve can be generated to evaluate the specificity and sensitivity of the diagnostic model. By moving a cut line vertically through the probability plot, the number of correctly classified spectra can be determined at each interval. The sensitivity and specificity at each interval can be plotted against each other and the AUC of the resultant curve can be calculated (102, 103). Based on conventional definitions of efficacy, the area under the generated ROC can be classified as excellent (AUC: 0.90-1), good (AUC: 0.8-0.9), fair (AUC: 0.7-0.8), poor (AUC: 0.6-0.7) and a fail (AUC: 0.50-0.60).
2.4 Performance of Pre-processing and Analytical Techniques using ex vivo Data Sets

2.4.1 Ex vivo Liver Data Sets

A collaboration with Dr. Kevin Hewitt’s lab at Dalhousie University (3) allowed us to obtain an ex vivo set of biological tissue. Mice livers with either normal or fatty characteristics were measured with the newly updated endoscopic Raman system. Frozen bulk liver lobes from adult male Rats in various stages of liver steatosis (fatty liver) were also measured, although with the old Raman System. The acquired Raman spectral peaks from the mice livers were identified and assigned to appropriate biological bond vibrations. Spectra from the rats were then classified using a Principle Component Analysis (PCA) approach. Differences between the old and new Raman systems were also seen.

Spectral analysis of Mice Raman spectra

Bulk ex vivo livers from Mice were measured with the newly updated Raman system (design discussed in Section 2.2). The mice were either control subjects fed a standard diet which resulted in normal livers, or were experimental subjects fed a high fat diet to induce a fatty liver. Raman spectra were measured on two normal mouse livers and three experimental fatty livers. Multiple measurements (5-10) were taken for each tissue sample and then averaged. A final average of each category (Normal or Fatty) was then produced.

The fluorescence of the raw spectra was removed through the Vancouver Algorithm (VA) as discussed in Section 2.3. Normalization was accomplished by dividing each
wavenumber by the area under the curve (AUC). The mean spectra of two normal mice livers and three fatty mice livers can be seen in Figure 2.6, where the difference between these two spectra is shown in Figure 2.7. The difference spectrum (Figure 2.5) shows peaks at 2846 cm\(^{-1}\), 2891 cm\(^{-1}\), and 3014 cm\(^{-1}\), respectively, corresponding to CH\(_2\) symmetric stretch, CH\(_2\) asymmetric stretch and =C-H stretch vibrations (85). At the same time there is a reduction in vibrations due to CH\(_3\) symmetric and asymmetric stretch (85).

\textbf{Figure 2.6:} Average Raman spectra of normal and fatty intact mice livers using the newly updated Raman system.
**Figure 2.7:** Difference spectra generated by subtraction of the control (Normal) mice spectra from the fatty mice spectra.

**Statistical Analysis of Rat Livers**

Bulk *ex vivo* livers from rats was also measured. However, these measurements were conducted using the old Raman system (design discussed in section 2.2). The rats were fed a
control diet or a high fat diet that was meant to induce a fatty liver. The livers were harvested at six different weeks on the high fat diet and Raman measurements were subsequently taken. Multiple measurements (3-10) from each liver were taken and averaged in order to account for tissue heterogeneity. It was found that removing underlying fluorescence from the samples using a polynomial fit led to artefacts in the processed Raman spectra due to its complex shape. To overcome this problem the underlying fluorescence was minimized on the larger rat data set by converting the calibrated and smoothed raw spectra into second order derivative spectra with a Savitzky-Golay six-point quadratic polynomial (discussed in detail in Appendix B). Normalization was accomplished by summing the squared derivative values of a spectrum and then dividing each variable in the spectrum by this sum. The complete data set of second order derivative spectra were then analyzed using statistical software (Statistica 6.0, StatSoft Inc. Tulsa, OK). Principal components (PCs) for all the spectra were computed to reduce the number of variables. This was followed by a linear discrimination analysis (LDA) with leave-one-out cross validation on single, or groups of, PCs that individually accounted for 0.1% or more of the variance.

Figure 2.8 shows the 1st principal component (PC1) contributions (PC1 coefficient × PC1 spectrum) to the total spectra over the six weeks. These spectra represent the amount of spectral change occurring compared to the overall average spectrum. The spectra look similar to lipid-like spectra, as expected if lipids are the principal source of the difference. Spectra from later weeks have more lipids than the average, earlier weeks have fewer lipids than the average. Spectra lying on the x-axis represent the average lipid signal for all six weeks. The spectra of earlier weeks fell below the x-axis therefore they have fewer lipids than the average and spectra from later weeks have more lipids than the average. The control is significantly lower than the
rest of the spectra. This measurement and assessment can be obtained in real-time (i.e. within a few seconds), on bulk tissue without preparation, which has clinical relevance for organ transplants (3). Moreover, it can be seen that the use of multivariate statistical techniques are able to clearly differentiate the livers based on lipid content.

Figure 2.8: The contribution of factor one converted back into a spectrum for each week on the fatty liver diet. Spectra falling above the x-axis (average spectrum) have higher lipid content than the average and spectra falling below the x-axis have lower lipid content than the average.
2.5 Development and Validation of *in vivo* Analytical Techniques

2.5.1 *In Vivo* Oral Cavity Data Set

Another data set, from collaborators in Dr. Kevin Guze’s lab at Harvard University (4), was sent to our lab for analysis. This set was *in vivo* Raman measurements collected in the oral cavity using the new Raman system as discussed in section 2.2.

The procedure conducted by our collaborators included eighteen treatment-naive subjects of both sexes between the ages of 20 and 65 years who were recruited from a panel of patients who were previously histologically diagnosed with either benign hyperkeratosis, inflammation, fungal, dysplasia (mild, moderate, and severe) or oral SCC at Dana Farber Cancer Institute and Brigham and Women’s Hospitals. The clinical presentation of the lesions studied varied. Patients were seated in a standard dental chair, and the tissue was dried with a gauze sponge around sites of interest. Then, Raman spectra were obtained in the emission range from 0 to 1800 cm\(^{-1}\) with the previously described Raman probe, using an acquisition time of 1 second. A specially designed probe cap was designed to maintain the probe at an optimum distance of 7.5 mm and perpendicular to the tissue surface, as well as to prevent ambient light from being collected during Raman measurements. Disposable plastic sleeves were placed over the probe as an infection control barrier. Spectra were recorded at 2 or more (2-9) different sites within a lesion, depending on size, and from the healthy tissue at a contralateral site. Again, the acquired Raman peaks were identified and analyzed and then spectra were classified into pathology groups using multivariate analysis.
Spectral Peak Analysis

The dark background was removed in real time from the raw spectra by using a spectrum acquired when the excitation laser was turned off. The spectra were then calibrated for the spectral response of the instrument using the known spectral irradiance of standard tungsten lamp (RS-10A, Gamma Scientific, San Diego, CA). The tissue autofluorescence as well as background fluorescence from the fibers (not completely filtered out) were removed from each spectrum through the VA (80). The spectra were then normalized by the area under the curve over the wavenumber range from 800 to 1700 cm\(^{-1}\). Once normalized, the spectra from the multiple sites within the same lesion were averaged in order to reduce the spatial variability within the lesion. One group, known as ≥MILD, contained spectra from all SCC (Squamous Cell Carcinoma) lesions and lesions with a high probability of developing into a malignant lesion (mild through to severe dysplasia). This group had 72 site measurements from 14 lesions. The second group, known as ≤LEU (Leukoplakia conditions and better), contained spectra from all benign lesions and from normal mucosa sites. This group had 46 site measurements from 22 lesions. Figure 2.9 shows two mean Raman spectra for all lesions within the ≥MILD (Mild dysplasia conditions and worse) and all lesions/normals within the ≤LEU groups respectively. Also shown in this plot is the difference spectrum between these two means. The mean spectra contained several prominent Raman peaks which were common to many human tissue types. The sharp peak centered around 958 cm\(^{-1}\) is consistent with a hydroxyapatite signal (85). The difference spectrum shows that this peak is more intense in the mean ≤LEU spectrum. The small peak at 1002 cm\(^{-1}\) is seen in both the ≥MILD and ≤LEU spectra at a similar intensity. This signal is caused by the symmetric ring breathing vibrational mode of the amino acid phenylalanine (104). There is a broad peak, ranging from roughly 1150 cm\(^{-1}\) to roughly 1250 cm\(^{-1}\). This peak is
due to multiple overlapping peaks at 1175 cm$^{-1}$ (C-H bending mode of tyrosine), 1204 cm$^{-1}$ (C-C$_6$H$_5$ stretching modes of tyrosine and phenylalanine (104)) and 1236 cm$^{-1}$ (amide III C-N stretching mode(104)). The signal at 1298 cm$^{-1}$ is due to a saturated fatty acid (105) and its intensity is greater in the ≤LEU spectrum than the ≥MILD spectrum. The small shoulder at 1337 cm$^{-1}$, which appears to have the same intensity in both mean spectra, can be attributed to the CH$_3$CH$_2$ wagging mode of collagen and emission from purine bases (104). The prominent peak at 1448 cm$^{-1}$ is found in most tissue types and is known to be due to CH$_2$ bending mode of proteins and lipids (104). The difference spectrum shows the 1448 cm$^{-1}$ peak is more intense for the ≤LEU group. The wavenumber range from roughly 1500 cm$^{-1}$ to roughly 1600 cm$^{-1}$ can be attributed to a large group of molecular vibrations, particularly protein ring assignments (104).

The slight shoulder found at 1618 cm$^{-1}$ was determined to be caused by the C=C stretching mode of tryptophan (104). It can also be seen that this shoulder is more prominent in the ≥MILD spectra. The relatively large peak found at 1650 cm$^{-1}$ is characteristic of amide I, C=O stretching mode of proteins which is indicative of an alpha helical conformation (104). The spectra also contain several small peaks of uncertain molecular origin. In addition there may be a small contribution to the spectra that can be attributed to emissions caused by the excitation light stimulating trace impurities in the optical fiber (106, 107).
Figure 2.9: Averaged ≥MILD (orange) and averaged ≤LEU (purple) Raman spectra after normalization by area under the curve, shifted along the y-axis for clarification. The black spectra shows the difference when the averaged ≥MILD is subtracted by the averaged ≤LEU.

Statistical Analysis

Multivariate techniques were used to compare the entire spectra, as opposed to single Raman peaks using the software STATISTICA 10.0 (StatSoft Inc., Tulsa, OK), in order to
determine the success of separating the spectra into the two pathology groups, \( \geq \text{MILD} \) and \( \leq \text{LEU} \). A set of principle components (PCs) were first calculated for all the spectra to reduce the number of variables (76). The PCs that accounted for >0.1\% or more in the variance were then analysed with Student’s t tests to determine which PCs were the most successful in separating the spectra into the two classification groups, \( \geq \text{MILD} \) or \( \leq \text{LEU} \). The most significant PCs were used in a linear discrimination analysis (LDA) with the leave one- out, cross-validation procedure. The number of PCs that were chosen for LDA was limited to 4 or less to avoid over-training of the data (70).

Using the data output from the cross-validation, an LDA posterior probability plot, and a Receiving Operator Characteristic (ROC) curve were generated. Figure 2.10 shows the posterior probabilities of each lesion or normal being classified as \( \geq \text{MILD} \) as generated through LDA with 4 principle components using a leave-one-out cross validation procedure. The symbols used for each lesion indicate the tissue pathology as determined by histology examination of subsequent biopsies. From the probability plot, an ROC curve was generated, shown in Figure 2.11, to evaluate the specificity and sensitivity of the diagnostic signal. Based on conventional definitions of efficacy, the area under the generated ROC of 0.916 was excellent. From the ROC curve it was determined that all the lesions with a pathology \( \geq \text{MILD} \) could be identified by the LDA model (100\% sensitivity) at the cost of identifying 5 false positives (77\% specificity). This has been summarized in Table 2.1 which shows the LDA classification results with a cut line in such a location that identifies all the \( \geq \text{MILD} \) lesions correctly.
Figure 2.10: Posterior Probability plot of an oral lesion being classified of ≥MILD as generated through LDA with 4 principle components using a leave-one-out cross validation procedure. The ≤LEU (benign) lesions are represented by diamond symbols and the ≥MILD (diseased) lesions are represented by square symbols.
Figure 2.11: Receiver Operator Characteristic (ROC) curve demonstrating the different specificities and sensitivities based on a cut line moving from 0 to 100% in Figure 2.10. The AUC of this ROC is 0.916.
Table 2.1: Summary of samples measured and LDA classification success constrained by the requirement of all lesions with a pathology ≥MILD have to be identified.

<table>
<thead>
<tr>
<th>Pathology</th>
<th>No. of Sites Measured</th>
<th>No. Of Lesions Measured</th>
<th>No. Of Lesions Correctly Classified</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>34</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td>Inflammation</td>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Fungal (Candida sp.)</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Leukoplakia</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mild Dysplasia</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Moderate Dysplasia</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Severe Dysplasia</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Squamous Cell Carcinoma (SCC)</td>
<td>57</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>118</strong></td>
<td><strong>36</strong></td>
<td><strong>31</strong></td>
</tr>
</tbody>
</table>

2.6 Chapter Conclusions

Two Raman systems and a specially constructed Raman probe have been designed and evaluated in section 2.2. Through the use of a Tylenol standard sample we have found that the newer Raman system which is constructed using the latest technology for the CCD and spectrograph has a higher signal to noise ratio.
Two *ex vivo* data sets (one from mice tissue and one from rat tissue) were measured using the new and old system respectively. Each data set was subjected to pre-processing techniques and it was determined that the Raman signals collected with the newer system were able to be extracted with the more reliable 5\textsuperscript{th} order polynomial. Data collected with the old system required the second order derivative to reduce the fluorescence background. However, PCA followed by an LDA was still able to show differences in the Raman spectra of the old system based on lipid content.

Finally, the *in vivo* oral cavity data set analysis demonstrates the potential for a clinical application of Raman spectroscopy to identify and differentiate premalignant/malignant oral lesions, and the overall potential for using Raman-based devices in a clinical setting. Using multivariate analysis it was demonstrated that the Raman spectra identified were able to discriminate between $\geq$MILD and $\leq$LEU with 100% sensitivity and 77% specificity. Despite the relatively small sample size, the LDA showed that the two groups of spectra were well separated in probability space.

Through optimization of hardware components, and analytical techniques, it was determined that Raman Spectroscopy was able to elucidate biochemical differences between different tissue pathologies. These results are promising for the use of Raman as a clinical tool. However, they are small sample sizes which need to be validated with larger studies. One such large study, which tests adjunct Raman Spectroscopy as a clinical tool in the lung, is presented in the next chapter.
Chapter 3: Real-time endoscopic Raman spectroscopy for *in vivo* early lung cancer detection in the Central Airways

3.1 Chapter Introduction

Raman spectroscopy is able to provide very specific, fingerprint-like spectral features in contrast to the broad spectral features of fluorescence and reflectance technology. Further, Raman spectroscopy has high accuracy for differentiation between malignant and benign tissues (62). It is potentially useful for improving the specificity of early lung cancer localization. However, the Raman signal is exceedingly low, and as a consequence, long integration times are required to acquire sufficiently strong Raman signals for a single spectrum. A traditional Fourier-transform Raman system requires up to 30 minutes of integration time to acquire one Raman spectrum with reasonably good signal-to-noise ratio (SNR). This has hindered the clinical applications of Raman technology. Our group has successfully developed a rapid, real-time Raman spectrometer system and a dedicated endoscopy Raman catheter for lung measurements that substantially reduces the spectral acquisition time to less than 1 second (77). This system employed proprietary technologies for improving the spectrometer’s SNR (108), Raman catheter fiber background fluorescence elimination, and size miniaturization (77). In a pilot study, we conducted point Raman spectroscopy measurements of suspicious areas identified by WLB+AFB imaging for improving early lung cancer detection (76). The results very well demonstrated the technical feasibility and clinical compatibility. Further improvements to the system as described in Chapter 2 were also implemented (2). This chapter reports a large scale single centre clinical study using the developed Raman spectroscopy technology as an adjunct device to WLB+AFB exam to improve the specificity of localizing malignant lesions (HGD/CIS) of the central lung airways, while maintaining high detection sensitivity.
3.2 Patients and Methods

This study was approved by the University of British Columbia – BC Cancer Agency Research Ethics Board (certificate number: H06-00010). Patients who were attending the BC Cancer Agency Vancouver Center for a previously scheduled bronchoscopy were invited to volunteer for this study. Patients must have already consented to a bronchoscopy as part of a standard diagnostic procedure or as part of an approved lung cancer prevention study, before being approached to volunteer. Patients were excluded if they had a cardiac pacemaker or implanted defibrillator device, had a known allergic reaction to Xylocaine, were taking a blood thinner such as warfarin or heparin, or had any medical condition such as acute or chronic respiratory failure, which could jeopardize the safety of the patient during participation in the study. Women who were premenopausal were excluded unless they were surgically sterile or on the birth control pill.

The Raman system used to take measurements was described in Chapter 2. Briefly, the system included a thermo-electrically cooled CCD detector for faster start up times and reduced optical noises (etaloning effect) (78), as well as a new spectrograph with a holographic reflection type grating, allowing tunable wavelength range (2). These changes were implemented to improve the SNR, as previously discussed in Chapter 2, thus allowing for a more reliable extraction of the Raman signal from the fluorescence background instead of the more obtuse and glossy 2nd order derivative processing of the data used by us previously (76). Figure 3.1 shows the schematic diagram of the endoscopic laser Raman spectroscopy system. The inserts show the arrangement of the excitation (red) and collection fibers (green). The Raman excitation light was produced by a wavelength stabilized 785 nm diode laser, and delivered to the tissue surface by a
detachable 1.8 mm size fiber optic probe (Raman catheter) passed down the instrument channel of the bronchoscope.

**Figure 3.1:** Schematic diagram of the endoscopic laser Raman Spectroscopy system. The inserts show the arrangement of the excitation (red) and collection fibers (green). The collection fibers were connected to the spectrograph through a special round to parabolic fiber bundle to correct the spectral imaging distortion to achieve better SNR.
The maximum excitation power at the tissue surface was 150 mW. The same catheter collected emission from the tissue and delivered it to the spectrometer for analyses. The collection fibers were connected to the spectrograph through a special round to parabolic fiber bundle to correct the spectral imaging distortion to achieve better SNR. The catheter itself contained ultra low OH impurity fibers, and a gold coated excitation fiber to avoid cross-talk between the excitation and collection fibers. It should be noted that the visible guidance channel, discussed in Chapter 2.2, was removed in this probe to allow for more collection fibers and thus a greater collection area. Optical filters were coated at the distal end of the probe to filter out laser noise, fiber emission, and to attenuate all collected light with wavelengths $\leq 820$ nm ($\leq 540$ cm$^{-1}$ relative to the 785 nm excitation). At its proximal end, the probe was attached to a second set of optical filters with similar transmission characteristics, but higher optical density in the rejection bands for further, and better, attenuation of the aforementioned unwanted emissions. The spectrograph grating was tuned to cover the high wavenumber range of 2050-3100 cm$^{-1}$ which is known to have much less tissue autofluorescence than lower wavenumbers ($< 1800$ cm$^{-1}$), and yet still have Raman bands sensitive to biomolecular changes (76, 77). The spectral resolution in this wavenumber range was estimated to be 8 cm$^{-1}$. Although the SNR increases with excitation energy ($power \times time$), an acquisition time of 1 second was used as this was the time that the catheter and excitation spot could be reliably maintained in the same position on the tissue surface. After data acquisition, custom designed software was used to subtract the fluorescence background using a modified polynomial fit ($9^{th}$ order) and to display the calibrated Raman spectrum all within a fraction of a second (80). An example of the fluorescence background removal is shown in Figure 3.2.
The procedure was for patients to first undergo a standard WLB+AFB exam which took place in the Endoscopy Suite and the location of any autofluorescence positive lesions identified. A three stage WLB+AFB visual grading of each site was determined by the physician following the criteria listed in Table 3.1 (37). After the AFB, the Raman probe was inserted into the instrument channel of the bronchoscope by the physician and directed toward an

**Figure 3.2:** An example showing 9th order polynomial fitting to the fluorescence background of the raw spectra of *in vivo* lung tissue, and the resultant Raman spectra.
autofluorescence positive lesion (a lesion identified through AFB and visually coded 2 or 3) or a normal tissue control site. A probe tip to tissue distance of between 5-10 mm was used which generated an excitation illumination spot diameter on the tissue surface of between 2 - 4 mm. A lesion was measured between 1-6 times, depending on size, and different lesions were considered to be separate sites with unique distinct pathology for analyses. Biopsies were taken as directed by the physician, and analyzed by an experienced pathologist using standard histopathology assessment for lung lesions (37). The pathological coding system and the corresponding tissue diagnosis are listed in Table 3.2. If different biopsy fragments of the same tissue site had different histological assessments, the worst assessment was taken for analysis. Multiple Raman spectra from a tissue site were averaged to generate a single spectrum to represent the entire site. The final data file matched the histopathology results to the WLB+AFB visual grades and the Raman spectrum for each site if all possible.
**Table 3.1**: Physician visual grading system and the corresponding tissue description.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Normal)</td>
<td>No visual abnormality</td>
</tr>
<tr>
<td>2 (Abnormal)</td>
<td>Visual changes suggestive of inflammation, trauma, hyperplasia, metaplasia, mild dysplasia</td>
</tr>
<tr>
<td>3 (Suspicious)</td>
<td>Visual changes suggestive of moderate dysplasia, severe dysplasia, carcinoma <em>in situ</em> or invasive cancer</td>
</tr>
</tbody>
</table>
Table 3.2: Pathological coding system and the corresponding tissue diagnosis.

<table>
<thead>
<tr>
<th>Pathological Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Normal</td>
</tr>
<tr>
<td>2</td>
<td>Inflammation</td>
</tr>
<tr>
<td>3</td>
<td>Hyperplasia or Metaplasia</td>
</tr>
<tr>
<td>4</td>
<td>Mild Dysplasia</td>
</tr>
<tr>
<td>5</td>
<td>Moderate or Severe Dysplasia</td>
</tr>
<tr>
<td>6</td>
<td>Carcinoma in Situ</td>
</tr>
<tr>
<td>7</td>
<td>Microinvasive Cancer</td>
</tr>
<tr>
<td>8</td>
<td>Invasive Cancer</td>
</tr>
</tbody>
</table>

Between May 2011 and November 2012, Raman spectra were obtained from 80 patients. Raman measurements were taken from a total of 280 sites. Of the 280 sites, 214 sites were biopsied after the Raman measurements and histologically assessed. Pathology reports showed that 72 sites were graded as either a HGD or worse, 90 sites were graded as benign (inflammation, metaplasia, hyperplasia or mild dysplasia), and 52 were graded normal. Although biopsying normal tissue was not standard surgical procedure, some normal control sites were biopsied as these were part of another study with a protocol which required the subject to consent to the taking of additional tissue. The remaining 66 control sites were not biopsied;
instead they were determined to be normal by the physician through visual grading during the WLB and AFB exam. Detailed patient demographics are shown in Table 3.3.

**Table 3.3:** Patient demographics and location of the lung Raman readings. A total of 280 sites were measured.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Patient Demographics</th>
<th>Raman Reading Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean age (range)</td>
<td>Male</td>
</tr>
<tr>
<td>Invasive SCC and CIS</td>
<td>66 (41-86)</td>
<td>21</td>
</tr>
<tr>
<td>Severe + Moderate Dysplasia</td>
<td>62(40-78)</td>
<td>32</td>
</tr>
<tr>
<td>Mild Dysplasia + Metaplasia</td>
<td>62(40-80)</td>
<td>53</td>
</tr>
<tr>
<td>Hyperplasia</td>
<td>67(40-78)</td>
<td>14</td>
</tr>
<tr>
<td>Inflammation</td>
<td>65(61-72)</td>
<td>6</td>
</tr>
<tr>
<td>Normal</td>
<td>67(40-84)</td>
<td>86</td>
</tr>
<tr>
<td>Total</td>
<td>211</td>
<td>69</td>
</tr>
</tbody>
</table>

### 3.3 Statistical Analysis

Multivariate statistical methods have been widely used for classification of Raman spectra in a number of studies related to cancer diagnoses including skin and were proven to be very effective and reliable (62). Introductions into pre-processing and mathematical modeling using PCA and PLS in relation to cancer diagnoses as performed by the author can be seen in Chapter 2 of this thesis. In this chapter, principal components with generalized discriminat
analysis (PC-GDA) and partial least-squares (PLS or called PLS-DA) were used for spectral classification with leave-one-out cross-validation (LOO-CV) where successive single spectra were left out for test with the remaining spectra used for training.

However, before the statistical analysis, all the Raman spectra were normalized to their respective integrated spectral areas under the curve (AUC). Analyses were performed on the full spectra based data set and also on data set with selected discrete wavebands generated using a number of feature selection strategies including: stepwise multiple regression (STEP) (93), least absolute shrinkage selection operator (LASSO) (95), and genetic algorithm (GA) (96). These feature selection algorithms have been previously utilized to improve model prediction/classification in a number of studies, for example, genomics and proteomics (109), lung cancer (88), breast cancer (89), and oral cancers (90). All the multivariate classification analyses in this study were implemented using MATLAB (version 2013b, Math-Works).

Waveband selection results depend on the sample spectra and sample size. In order to find reliable optimal wavebands, a LOO-CV protocol was used. In the LOO-CV waveband selection procedure, a single spectrum was left out with the remaining spectra used for waveband selection operation. A set of wavebands were selected which gave the best diagnostic performance of the training spectra. By repeating this procedure, every spectrum was left out once for wavelength selection purpose. At the end, n sets of wavebands were selected, where n was the total number of cases. The n sets of wavebands were then accumulated. The wavebands with higher odds from the LOO-CV analysis were chosen for subsequent PCA-DA and PLS analyses. Waveband selection was also tested using three-fold cross-validation where the spectra were divided equally and randomly into three groups with two groups for training and one group for testing. Similar results were obtained for three-fold cross-validation with those of LOO-CV.
In order to prevent over-fitting or selecting spurious wavebands, a horizontal window size of 5 pixels (successive horizontal binning of 5 pixels), corresponding to an average of 5.3 cm$^{-1}$ spectral range, was chosen based on recommendations from literature (91, 94, 110).

Following the pre-processing of spectra PC-GDA analysis was used following the methodology found in (99). Briefly, the training process consisted of the following procedures: (1) the mean and standard deviation of the training spectra data set were calculated, (2) each spectrum in the training data set was standardized by subtracting the mean and then dividing by the standard deviation, (3) the standardized training spectra were analyzed with principal component analysis. PC factors of the training cases and PC loadings were obtained, and (4) a generalized (linear) discrimination model was developed from the PC factors which could be used directly to predict the new cases. The testing process consisted of the following procedures: (1) the test spectrum was standardized by removing the mean and dividing by the standard deviation obtained from the training spectra, (2) the PC factors of the test spectrum was calculated based on the PC loadings from the training spectra, and (3) a posterior probability of the testing spectrum was obtained from the discrimination model developed in the training procedure. The above procedures were repeated until all the spectra were left out once (and only once) for testing.

For PLS analysis, the methodology found in (99) was used. The training process consisted of the following procedures: (1) the mean and standard deviation of the training spectra data set were calculated, (2) each spectrum in the training data set was standardized by subtracting the mean and then dividing by the standard deviation, (3) the standardized training spectra were analyzed with NIPALS algorithm (nonlinear iterative partial least squares) with the classification of the training spectra setting to the known value. The weight factors, the loadings,
the regression coefficient and the factor scores of the training spectra were obtained, (4) a
general discrimination model was developed from the training spectra which could be used
directly to predict new cases. The testing process consisted of the following procedures: (1) the
test spectrum was standardized by removing the mean and dividing by the standard deviation
obtained from the training spectra, (2) the factor scores of the test spectrum were calculated
based on the weight factors from the training spectra, and (3) A posterior probability of the test
spectrum was obtained from the discrimination model developed in the training procedure. The
above procedures were repeated until all the spectra were left out once (and only once) for
testing.

The receiver operating characteristic (ROC) curve was calculated from the posterior
probabilities derived from each of the analysis models described above and represents the
diagnostic performance of each model. The AUC of each ROC was calculated using the
trapezoidal rule (102). The significance of these AUCs and comparisons between different AUCs
were carried out in a standard fashion (103). All ROC analyses were based on nonparametric
techniques and were conducted separately for the PC-GDA and PLS analyses. To compare the
different statistical methods used, and to compare the utility of Raman spectroscopy with other
non-invasive diagnostic techniques, the specificities were calculated for fixed sensitivity levels
of: 90% and 95%.
3.4 Results and Discussion

3.4.1 Mean Raman Spectra Peak Assignments

The Raman spectra were dominated by strong CH stretching related bands in the wavenumber range from 2775 to 3040 cm\(^{-1}\). For wavenumbers below 2775 cm\(^{-1}\) the spectra did not contain any clear Raman signals, but a significant amount of noise attributed to etaloning effects and residual optical fiber emissions. The mean Raman spectra from 2775 to 3040 cm\(^{-1}\) for each histopathology group are shown in Figure 3.3. Mild dysplasia and metaplasia (pathology categories marked as 3 or 4) were grouped together because both pathologies are considered to be low grade preneoplastic lesions, carrying the same risk of progression. Similarly, moderate and severe dysplasias were grouped together because both are considered to be high grade preneoplastic lesions which carry a similar risk of progression into invasive cancer. Furthermore, due to a small number of CIS lesions (n=2) as well as the higher risk of progression into invasive cancer, CIS was grouped with invasive cancers (IC). No micro-invasive cancer cases were obtained in this chapter. For these reasons, Figure 3.3 only has 6 categories. Major Raman peaks are seen at 2850, 2885, 2940, 2965, 2990, and 3020 cm\(^{-1}\). These peaks were assigned to various fundamental CH, CH\(_2\), and CH\(_3\) stretching modes (85) and overtones of CH\(_2\) and CH\(_3\) bending modes (111). The peak at 2850 was assigned to the CH\(_2\) symmetric stretching modes of fatty acids and lipids, while the peak at 2885 cm\(^{-1}\) was for the CH\(_3\) symmetric stretching modes (85).
Figure 3.3: Mean Raman spectra by diagnosis. All spectra were normalized to their respective area under curve (AUC) before averaging by diagnosis. CIS: carcinoma in situ.

The main peak at 2940 cm\(^{-1}\) was assigned to a mixture of CH vibrations in proteins and CH\(_3\) asymmetric stretching modes of lipids and nucleic acids (76, 85). And the peaks at 2965 and 2990 cm\(^{-1}\) were assigned to in-plane and out-of-plane anti-symmetric CH\(_3\) stretching in lipid and fatty acid molecules (85, 105, 112). The peak 3020 cm\(^{-1}\) was assigned to the asymmetric stretching of =C-H group in RCH=CHR molecules (112), in which the R stands for an alkene functional group.

Apart from the main Raman peaks there was evidence for smaller Raman peaks, or inflection points at 2790, 2825 and 2920 cm\(^{-1}\) that did not appear to be related to noise. The
origins of these bands were uncertain. The 2920 cm\(^{-1}\) band was most likely due to Fermi resonance interactions between the main stretching modes and CH bending overtones. Whereas the 2825 cm\(^{-1}\) band may be due to one of the pair of CH stretching modes of aldehydic functional groups, with the other lost in the noise at lower wavenumbers (113). No explanation can be offered currently for the 2790 cm\(^{-1}\) band.

Despite there being clear Raman bands that were probably connected to the abundance of different biomolecules, there were no unique peaks that could be assigned to lung cancer alone yet, there were some specific characteristics in the main Raman peaks that correlated with histopathology grades. On average there was a distinctive loss of the lipid peak at 2850 cm\(^{-1}\) seen in the spectra from malignant lesions, the amount lost for individual lesions showed considerable variation. For example, IC had spectral shapes that consistently demonstrated a reduced intensity of the CH\(_2\) symmetric stretch emissions at 2850 cm\(^{-1}\) and increased intensity at 2940 cm\(^{-1}\) from CH\(_3\) asymmetric stretching modes. Movasagi \textit{et al} suggested that the 2850 cm\(^{-1}\) was a good indicator for the change in the amount of lipid in the samples (85). The 2940 cm\(^{-1}\) peak has been associated with both proteins and nucleic acids (76, 85). A previous study that used Raman spectroscopy on \textit{ex-vivo} lung tissue samples also found a reduction in the lipid content for malignant tumor tissue (114). This was attributed to the decrease of phospholipids found in the cancerous tissue when compared to normal controls. The same study also found that there was an increase in certain amino acids while others decreased, and many of these are known to have strong Raman emissions in the 2775 to 3040 cm\(^{-1}\) range (85). Thus changes in the abundance of these amino acids may have contributed to the changes seen in the Raman spectra measured in this study.
There were also differences in the spectra from inflamed tissue compared to the other pathologies. Specifically, intensity of the inflammation group was relatively higher than all other categories between 2850 and 2900 cm\(^{-1}\). Although there were only a small number of spectra from inflamed tissue, all six cases came from six different patients, indicating that the abnormalities in the inflammation spectra were most likely not due to chance. Inflammation has been suggested to be a risk factor in certain cancers (115), particularly chronic inflammation. All six biopsies from the inflammation sites contained both chronic and acute inflammation histopathology, meaning that the Raman spectra could not be separated into the two subgroups. Due to this co-pathology, it remains unknown if the difference in the Raman readings from inflamed tissue was due to the chronic or acute diagnosis. Nevertheless the unique Raman signatures of inflammation tissues support the idea that Raman spectroscopy can help to improve the diagnostic specificity of AFB because inflammation is a significant cause of false positives in AFB lung cancer localization (116).

### 3.4.2 Multivariate Statistical Analysis

To extract a more reliable correlation of spectra with pathology, multivariate statistical techniques were used as described in section 3.3. The results from multivariate statistical analysis demonstrated that premalignant and malignant lung lesions can be differentiated from benign lesions and normal tissue using point laser Raman spectroscopy. Our lab has previously shown that this was possible using a model algorithm based on linear discrimination analysis (LDA) of spectra processed using a 2\(^{nd}\) order derivative to remove the background fluorescence (76). The findings in the current study were obtained for a much larger sample size, using a
system with alternative components to reduce system noise, and focus solely on the pure Raman spectra, rather than their noisier 2nd order derivatives, as discussed in Chapter 2.

The spectral range from 2775 to 3040 cm\(^{-1}\) was chosen as the optimal wavenumber region for analyses as it contained the only significant Raman emissions in the range from 2050 to 3100 cm\(^{-1}\) which was measured. These emissions were assigned to various CH stretching modes, although a small peak at 2790 cm\(^{-1}\), which helped with the discrimination between groups, was unidentified. The optimal upper bound cut off was determined to be 3040 cm\(^{-1}\) as the signal beyond this wavenumber had significantly more noise. Although Raman emissions due to water molecule stretching modes that occur for wavenumbers from 3200-3500 cm\(^{-1}\) were outside the range of our system.

As the presence (or absence) of a single wavenumber or peak could not be attributed to lung cancer alone, various analysis were conducted and can be seen in table 3.4 and 3.5. Different groupings were explored in order to find the optimal algorithm for a classification model. An optimal classification model – a model will high predictive accuracy - will group members of the same pathology together.
Table 3.4: Area under the ROC curve and 95% CI based on full range and waveband selection algorithms for discriminating premalignant (HGD) and malignant lung cancers (n = 72) from benign lung diseases and normal lung tissues (n = 208).

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Full range</th>
<th>STEP</th>
<th>LASSO</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLS analysis</td>
<td>0.83</td>
<td>0.88</td>
<td>0.86</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>(0.77-0.88)</td>
<td>(0.84-0.92)</td>
<td>(0.81-0.91)</td>
<td>(0.80-0.90)</td>
</tr>
<tr>
<td>PC-GDA analysis</td>
<td>0.83</td>
<td>0.88</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>(0.78-0.88)</td>
<td>(0.84-0.92)</td>
<td>(0.80-0.90)</td>
<td>(0.80-0.90)</td>
</tr>
</tbody>
</table>
**Table 3.5:** Summary of Raman spectroscopy diagnostic parameters derived from ROCs: specificity values according to different levels of sensitivity for full range and waveband selection algorithms. HGD and malignant (n = 72) versus benign lung lesions and normal lung tissues (n = 208).

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity level (95% CI)</th>
<th>Full range</th>
<th>STEP</th>
<th>LASSO</th>
<th>GA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PLS analysis</strong></td>
<td>0.95 (0.88-0.99)</td>
<td>0.42 (0.35-0.49)</td>
<td>0.49 (0.42-0.56)</td>
<td>0.46 (0.39-0.53)</td>
<td>0.44 (0.37-0.51)</td>
</tr>
<tr>
<td></td>
<td>0.90 (0.81-0.96)</td>
<td>0.51 (0.44-0.58)</td>
<td>0.65 (0.58-0.71)</td>
<td>0.53 (0.46-0.60)</td>
<td>0.53 (0.46-0.60)</td>
</tr>
<tr>
<td><strong>PC-GDA analysis</strong></td>
<td>0.95 (0.88-0.99)</td>
<td>0.40 (0.35-0.49)</td>
<td>0.46 (0.39-0.53)</td>
<td>0.43 (0.39-0.53)</td>
<td>0.44 (0.37-0.51)</td>
</tr>
<tr>
<td></td>
<td>0.90 (0.81-0.96)</td>
<td>0.51 (0.44-0.58)</td>
<td>0.64 (0.57-0.70)</td>
<td>0.46 (0.39-0.53)</td>
<td>0.50 (0.43-0.57)</td>
</tr>
</tbody>
</table>
When the Raman spectra (full range) were used to distinguish premalignant (moderate and severe dysplasia) and malignant (CIS & invasive cancers) lung tissues (total n=72) from benign (mild dysplasia, metaplasia, hyperplasia, and inflammation) and normal lung tissues (total n=208), the AUCs of the resulting ROCs for PLS and PC-GDA analyses were almost identical at 0.83 (column 2, Table 3.4) and statistically significant (p<0.001). Waveband selection techniques STEP, LASSO and GA were applied before PC-GDA and PLS analysis to improve the diagnostic performances. Figure 3.4 and Figure 3.5 show the wavebands selected by LASSO and STEP respectively. Some of the selected wavebands, but not all of them, were located at the Raman peaks. Some regions off the peak positions of the spectra were also selected and able to help with classification as well. The three waveband selection methods increased the AUC of their respective ROCs to between 0.85 and 0.88 (columns 3 - 5, Table 3.4). The diagnostic power of these analysis methods are shown in Table 3.5. Although the addition of STEP as a feature selection gave the best classification performance, the two other waveband selection methods (LASSO or GA) returned similar AUC values, indicating the discrimination results were reliable.

The sensitivities and specificities generated from these ROC curves were high as shown in table 3.5. At 90% sensitivity, 65% specificity is achieved by the STEP-PLS method in localizing HGD and malignant lesions. In comparison the AFB clinical trial that led to its FDA approval had a similar specificity of 66%, but the sensitivity is 67% only (37). At a sensitivity level of 90%, the full spectrum PLS analysis and PC-GDA analysis provided the same specificity of 51% (95% CI: 0.44-0.58). Use of the waveband selection techniques STEP, LASSO and GA increased the diagnostic specificity (columns 4 - 6, Table 3.5). For a 90% (95% CI: 0.81-0.96) sensitivity the PLS analyses with STEP waveband section provided the best
specificity of 65% (95% CI: 0.58-0.71). As an example, the posterior probabilities and the ROC curve corresponding to the STEP PC-GDA analysis are shown in Figure 3.6.

**Figure 3.4:** Mean Raman spectra by diagnosis, with highlighted areas showing the selected wavebands by LASSO. All spectra were normalized to their respective area under curve (AUC) before averaging by diagnosis.
Figure 3.5: Mean Raman spectra by diagnosis, with highlighted areas showing the selected wavebands by Stepwise. All spectra were normalized to their respective area under curve (AUC) before averaging by diagnosis.
Figure 3.6: Lesion classification by Raman spectroscopy based on STEP PC-GDA analysis.

(A) The posterior probability plot for distinguishing cancerous lesions (HGD, CIS and invasive cancer; n=72) from benign lesions (mild dysplasia, metaplasia, hyperplasia, inflammation and normal; n=208).  

(B) The ROC curves and 95% CIs derived from the posterior probabilities.
Figure 3.7: Lesion classification by Raman spectroscopy based on STEP PC-GDA analysis (box plot). The box plot representation of the posterior probability distributions according to lesion subcategories. The lower bound on each box shows the 25th percentile, where the upper bound shows the 75th percentile, meaning 25% and 75% of the data points are below these two bounds respectively. The line between these two bounds is the median. The whiskers on the plot show the 5th and 95th percentile, meaning 5% and 95% of the data are below these whiskers respectively. Data found outside the 5-95th percentile whiskers are outliers, as shown by separate data points.
A trend for the spectra as a whole was apparent as shown in Figure 3.7. A relatively large change in the posterior probability occurs between the mild dysplasia and the moderate dysplasia, indicated by the differences between the groups in the box plot. This high probability trend continues for Raman spectra from malignant tissue sites. All four multivariate techniques showed this trend. These would appear to indicate that there was a significant change in the spectra from tissue with HGD compared to those from benign or normal tissue, and this change becomes more pronounced for malignant tissue. This is consistent with increasing severity of bio-molecular changes that accompany HGD and malignancy transformations respectively.

Currently it is known that not every HGD will progress into a malignant state (15), and those which progress are not able to be visually determined by AFB or WLB endoscopy from those which do not. Understanding the biochemical signatures of those lesions which spontaneously regress versus those which do not, would provide valuable clinical information. Early detection, especially detection of lesions which would be known to become invasive could improve therapy.

**Classification based on modeling malignant versus normal lung tissues**

A different classification of all the histopathologies (n = 72 vs n = 208) was tried by using models generated from spectra of the two extremes only: IC (n=29) and normal sites (n=118). These models were then used to classify all histopathologies. The rationale for this was that all histopathologies were expected to lie between IC and normal sites according to their probability to be a cancer. These tests provided an ROC AUC of 0.85 (95 % CI: 0.81-0.90), based on full spectrum PC-GDA, and an ROC AUC of 0.86 (95 % CI: 0.81-0.91) based on full spectrum PLS. The posterior probabilities and the ROC curve corresponding to the full spectrum
PC-GDA analysis are shown in Figures 3.8 and 3.9, and demonstrated that all the histopathologies, including those categories not trained, can be well classified based on this approach. Similar trends as Figure 3.7 were found in Figure 3.9.

The trends in Figure 3.7 and Figure 3.9 suggest a hypothesis that the posterior probability values calculated from Raman spectrum may be an indicator of the likelihood of a HGD lesion progressing to malignancy. The posterior probability values may also be used to predict the prognosis of invasive cancers. These hypotheses are worth to be tested in future studies.
**Figure 3.8:** Lesion classification based on modeling using the extremes cases (29 Invasive Cancers and 118 Normal Sites) for training (full spectra PC-GDA). Testing was done for the entire data set (72 HGD/malignant vs. 208 benign/normal). (A) The posterior probability plot for distinguishing cancerous lesions (HGD, CIS and invasive cancer; n=72) from benign lesions (mild dysplasia, metaplasia, hyperplasia, inflammation and normal; n=208). (B) The ROC curves and 95% CIs derived from the posterior probabilities.
**Figure 3.9:** Lesion classification based on modeling using the extremes cases (29 Invasive Cancers and 118 Normal Sites) for training (full spectra PC-GDA) (box plot). Testing was done for the entire data set (72 HGD/malignant vs. 208 benign/normal). The box plot representation of the posterior probability distributions according to lesion subcategories.
3.4.3 Sensitivity and Specificity of Visual Diagnosis

Before performing Raman spectrum measurement, a visual grade was also determined for the majority of sites during the WLB+AFB exam by the physician (Table 3.1). There were three categories of visual grading: grade 1 for normal, grade 2 for abnormal (suggestive of inflammation, metaplasia, hyperplasia, and mild dysplasia), and grade 3 for suspicious (suggestive of moderate dysplasia, severe dysplasia, CIS and IC cancers) (37). The total number of sites that had a WLB+AFB visual grade, a Raman spectrum, and a matching histopathology assessment from a biopsy was 193, of these 38 were positive for HGD and 24 were positive for IC/CIS based on histopathology as shown in Figure 3.10. The STEP-PLS diagnostic algorithms based on Raman spectra of these 193 sites has an area under the ROC curve of 0.85, very close to the value, 0.88 for the STEP-PLS algorithm generated from the full 280 tissue sites. When comparing the WLB+AFB visual grades to the corresponding histopathology, WLB+AFB grade 3 (suspicious) identified 27 out of 62 lesions that were moderate dysplasia or worse, representing a 44% sensitivity; the specificity was 68% because it correctly identified 89 of the 131 lesions that were mild dysplasia or lower pathology grade according to Figure 3.10. Alternatively visual grades 2 and 3 combined (suspicious + abnormal) identified 59 out of 62 lesions that were moderate dysplasia or worse, a sensitivity of 95%; the specificity decreased to 13 % because only 17 of the 131 benign lesions/normal sites were correctly identified (Figure 3.10). The posterior probabilities generated from the STEP-PLS Raman algorithm identified 24 more HGD and malignant lesions (17 HGD, 1 CIS, and 6 IC) than the WLB+AFB grade 3 while keeping the same specificity level of 68%. Thus the relative sensitivity improvement of WLB+AFB+Raman versus WLB+AFB grade 3 was \((27+24)/27 = 1.89\).
**Figure 3.10:** The histopathology distribution of bronchoscopy (WLB+AFB) visual grading of lung lesions and normal tissue sites. The number in the parenthesis is the total number of cases within a particular visual grading subcategory/histopathology subcategory combination.
Alternatively if we keep the same sensitivity of 95% provided by the WLB+AFB grades 2+3 combined, the same Raman algorithm (STEP-PLS) were able to correctly identify 48 more benign lesions/normal sites (18 normal, 1 inflamed, 13 metaplasia/hyperplasia, and 16 mild dysplasia) than WLB+AFB grades 2+3 combined. Thus the relative specificity improvement of WLB+AFB+Raman versus WLB+AFB grades 2+3 was (17+48)/17 = 3.82.

Although the adjunct use of AFB to WLB has significantly increased the detection of HGD compared to WLB alone, there remains the inherent problem of poor specificities (43). The specificity of the AFB visual grading used in this study ranged from 13% for grades 2+3 to 68% for grade 3 only. The corresponding sensitivities swung drastically from 95% to 44% respectively. Although the latter was a little lower than the values obtained in some AFB studies for similar specificities, it was not inconsistent with the high variability in sensitivities for high specificities as shown in the meta-analyses of Sun et al. (43) This variability highlights how difficult it is to consistently obtain high sensitivities with good specificities. The adjunct use of Raman spectroscopy made a significant improvement to the specificity of detecting lesions in this study with HGD or worse. If high specificity is required, it was shown that for keeping the 68% specificity achieved by WLB+AFB visual grade 3, Raman identified 89% more true positives (51 sites compared to 27 respectively), representing a 1.89 times improvement in sensitivity. Alternatively if high sensitivity is required, it was shown that for keeping 95% sensitivity achieved by WLB+AFB visual grades 2+3 combined, Raman reduced the false positives by 42% (66 sites compared to 114 respectively), representing a 3.82 times improvement in specificity.

This chapter showed that the adjunct use of Raman spectroscopy improves the specificity of detecting premalignant or malignant lung lesions of the main airways compared to AFB visual
grading. No loss in sensitivity by Raman diagnosis occurred relative to WLB+AFB visual grades 2 and 3 combined; but the specificity increased significantly. Other methods have been tried to improve specificity, but with limited success. The fluorescence R/G ratio method looks particularly promising but so far only one study showed any significant benefit (117), but its performance at high sensitivity settings are still inferior compared to our Raman results: at 95% sensitivity Raman has a better specificity of 49%, compared to the 32% for the fluorescence R/G ratio method; at 90% sensitivity Raman has a better specificity of 65%, compared to the 53% for the fluorescence R/G ratio. These results lend support to the concept that Raman spectroscopy provides the physician with objective secondary information compared to the more subjective visual appearance of lesions using WLB/AFB examination. This will reduce the number of false positive biopsies, procedural time, and healthcare costs. Most importantly, maintaining high sensitivity (≥90%) is critical for accurate diagnosis and management.

Raman scattering within the main lung airways can be measured within 1 second and used to improving the localization of lung cancer and precancerous lesions. We envision that an algorithm derived from a database of Raman spectra would be able to classify a lesion in less than half a second, making this approach feasible for real-time lung cancer localization. Different from subjective interpretation of WLB/AFB images, the Raman algorithm represents an automatic, objective diagnosis based on quantitative Raman spectral analysis. Point Raman measurement on lesions identified by WLB+AFB visual grading 2 and 3 combined is a promising new clinical method for real-time localization of lung cancer/precancerous lesions.
3.5 Chapter Conclusions

A single centre clinical investigation of the adjunct use of real time Raman spectroscopy to the standard WLB and AFB for *in vivo* lung cancer localization of the central airways was conducted. *In vivo* real-time point laser Raman spectroscopy was performed on 280 lung lesions and normal tissue sites with a measurement time of 1 second per spectrum. Using multivariate techniques and waveband selection methods on the Raman spectra, it was shown that HGD and malignant lung lesions can be differentiated from benign lung lesions and normal lung tissues with high sensitivity (90%) and good specificity (65%). Compared to WLB+AFB visual grade 3 based diagnosis, Raman+WLB+AFB improved the sensitivity of localizing premalignant and malignant lesions by 1.89 times, while compared to WLB+AFB visual grade 2+3 combined diagnosis, Raman+WLB+AFB improved the specificity of localizing premalignant and malignant lesions by 3.82 times. Different from subjective interpretation of WLB/AFB images, the Raman algorithm could potentially become an automatic and objective diagnosis method based on quantitative spectral analysis. Further multi-center clinical trials are warranted to fully test the potential of this technology for improving lung cancer and precancerous lesion localization.
Chapter 4: Design and Development of an *in vivo* Peripheral Lung Raman Catheter

4.1 Chapter Introduction

In 2011 the National Lung Cancer Screening Trial was published in the New England Journal of Medicine (9). The study concluded that the addition of low-dose CT (LDCT) as a screening method for high risk lung cancer patients was able to reduce mortality rates when compared to the standard of the time, chest radiography (9). CT is an imaging technique with high sensitivity (around 89%) when compared to chest radiography (around 78%) (35). Therefore small lung nodules are found earlier and at a less invasive stage leading to more efficient therapeutic options for patients. As such, LDCT has become widely used clinically for the screening of lung cancer. As a result, there has been an increase in small nodules (<15mm) found in the peripheral airways.

However, it was also found that LDCT, when used specifically for lung cancer screening, has a false positive rate comparable to chest radiographs (9). The study published that over 95% of positive results found using LDCT were in fact false positives for lung cancer, suggesting that the vast majority of lung nodules found are in fact benign. Depending on size, location, further imaging and follow ups, some of these lesions are not found to be benign until it is biopsied and sent for histopathology.

Yet, there are significantly less options in the peripheral airways for imaging and localization of lung nodules. As standard white light and autofluorescent bronchoscopes can only reach the sixth division of the lung with reliability imaging extensions are used. Endo-bronchial Radial Ultrasound, discussed in detail in section 1.2.4, is often used to locate the nodule for biopsy, but the procedure has a low diagnostic yield and is currently unable to guide
physicians to a target lesion. 3D virtual bronchoscopy, a method that virtually reconstructs a patient’s bronchial tree from information gathered in the CT-scan, has been used to address this limitation (53). However, results are mixed and the technique requires extra trained and experienced personal (53, 118, 119). Further, the R-EBUS catheter uses the same instrument channel as the biopsy forceps, meaning the procedures cannot occur concurrently. Thus, biopsies are taken without direct visualization, with only the guidance of the R-EBUS sheath locked in place to guide the biopsy forceps to the tissue surface. This sheath can move out of place so multiple biopsies are usually taken to compensate for this possibility, increasing a patient’s risk for adverse events. Furthermore, the radial ultrasound probe is side-viewing, while the biopsy forceps are end-viewing, potentially causing biopsies to be retrieved from areas not found by R-EBUS.

There is a need to develop clinical technologies for this region of the lung which are capable of providing more information in real time and are well matched to the current clinical procedure. Previous Chapters have discussed the utility of using Raman spectroscopy as an adjunct diagnostic device in the central airways. This Chapter presents the design and development of a novel Raman catheter for the peripheral lungs.

4.2 Challenges for Peripheral Devices and Design Rationale

The most significant challenge in designing a new Raman catheter for peripheral airways is the architecture of the lung. The airway entrance to the lung, the trachea, first divides at the main carina, splitting into the right and left lung (14). Further divisions of the airways occur separating the right lung into upper (RUL, right upper lung), middle (RML) and lower lobes
(RLL) (14). The left lung is further divided into only upper (LUL) and lower lobes (LLL) – as the heart takes the space of where a middle lobe would be. The location of the heart also causes this first division of the trachea to be asymmetrical (120). The right main bronchus has a larger diameter than the left as well as a smaller angle with the longitudinal axis of the trachea (120). Further divisions into the right and left upper lobes also have a large angle away from the trachea – compounding the geometric restrictions on designing new lung catheters.

Specifically in peripheral lung cancer applications, the upper lungs are highly relevant. Lung Cancer of smokers predominately occurs in the upper lungs so technologies need to be able to navigate these sharp turns inside the airways (121, 122). Flexible catheters are going to be more successful at navigating this lung architecture than rigid ones, especially at the distal tip. A long rigid portion will not advance through a sharp turn. This requirement restricts the use of optical components at the probe tip. Conversely, there is the need for at least some overall rigidity because the probe needs to advance when pushed at the proximal end.

Another physical restriction encountered in the design of a new peripheral Raman probe is the need to use existing clinical standards. As the R-EBUS procedure locates a peripheral lesion and then locks a guide sheath in place to direct biopsy forceps, a new Raman catheter would need to go down this guide sheath as well. The guide sheath used at the BC Cancer Agency Vancouver Centre is the Single Use Guide Sheath Kit (Model K-201 KXK, Olympus Medical Systems, Tokyo, Japan). The inner diameter of the guide sheath is 1.4 mm, which is significantly smaller than the diameter of the central lung probe (1.8 mm). As well, the sheath is constructed of plastic, which imposes restrictions on friction. Two high friction plastics may not slide past one another easily, slowing down or stopping the advancement of the Raman catheter to the periphery.
Finally the phenomenon of Raman scattering events also pose a high demand on catheter design. As Raman is inherently weak, often to increase collection efficiency one must increase the collection area for a greater retrieval of scattered photons. Therefore, the greater the number of collection fibres in the given diameter the higher the collection efficiency.

4.3 Developmental Procedure

A probe was initially designed with the aforementioned challenges in mind, and based off the previously described central lung probe (5). Figure 4.1 shows the first generation peripheral lung probe. A 1.35 mm diameter fibre bundle was created with one central 100 µm excitation fibre and 31 105 µm collection fibres arranged in surrounding circles, as seen in Figure 4.2. The central fibre is aluminum coated (Model: AFS100/140/200A, Fiberguide Industries, Sterling, New Jersey) to prevent cross-talk with the collection fibres, which are made of low-OH silica (Model: AFS105/125/145T, Fiberguide Industries, Sterling, New Jersey). This bundle is encased in a sheath made of FEP (Fluorinated Ethylene Propylene) at the distal end (Zeus Industrial Products, Inc, Orangeburg, South Carolina), and encased in a stainless steel jacket (Zeus Industrial Products, Inc, Orangeburg, South Carolina) at the proximal end. The proximal end is separated so that the excitation and collection fibres can be properly coupled to the excitation source or detection path. This end is coupled using FC connectors for the ease of proper system alignment.
Figure 4.1: Specifications for the first generation peripheral lung Raman Probe.
4.3.1 Ferrule Removal

However, during the testing of the first generation probe down a standard bronchoscope and guide sheath it was observed that the probe could not advance through the maximum bend of the endoscope. The diameter challenge and the collection efficiency were addressed, so it was concluded that the rigid portion at the distal end was too long due to the metal ferrule and therefore needed to be removed. However, the metal ferrule was crucial in the alignment of the fibre bundle. Aligning the excitation fibre in the center of the bundle allows for the maximum collection of light scattered back and the metal ferrule facilitates the gathering of the collecting fibres and maintains the fibre bundle shape – essentially making this alignment much easier.

Yet, the need to advance past the sharp turn in the bronchoscope outweighed the benefits of the metal ferrule, so it was removed. As optical fibres are fragile, simply cleaving the distal end of the probe may result in the fibres cracking parallel to the axis, reducing the collection

Figure 4.2: End on view of the distal end of the first generation peripheral lung Raman probe.
efficiency. Thus, the metal ferrule was removed through extensive polishing with a coarse sand paper.

4.3.2 Etching Procedure

Another discovery during the testing of the first generation probe was the inability of the FEP material (which encases the distal part of the probe) to form a bond with epoxy. The FEP material was chosen for its low friction properties and flexibility. The low friction ensured a path through the bronchoscope without resistance, where the flexibility was necessary for the sharp turns. However, as seen in Figure 4.3, these characteristics also meant that it was unable to be bonded or create a water-tight seal.
Figure 4.3: Side view of the distal end of the peripheral lung Raman probe (prior to the metal ferrule removal). The gap is due to the FEP material characteristics showing that it is unable to bond with epoxy. This gap exposes the fragile optical fibres of the peripheral lung Raman probe.

Yet, the need to create a water-tight seal cannot be understated. Raman probes and catheters for \textit{in vivo} tissue applications are often designed to be re-useable as the optical components to make them are costly. Similarly, the peripheral lung Raman probe was also re-useable. This means that a single probe will be used \textit{in vivo}, disinfected and reprocessed, then used on another patient. Cracks and gaps not only expose the fragile fibres to a harsh cleaning process, but also create areas in which bacteria could hide and evade the reprocessing procedure. As such, it is necessary to maintain a sealed instrument.
A chemical etching procedure was therefore used to make the FEP bondable. The procedure removes the fluorine atoms at the surface of the FEP and forms a carbonaceous layer on the material that is compatible with adhesive. Once the metal ferrule was fully removed, the distal probe tip (5 mm) was etched using FluoroEtch Safety Solvent (Acton Technologies, Pennsylvania 18640 USA). The recommend procedure is to start by immersing the material in the solution for 30-60 seconds, while waving the material through the liquid to ensure a flow of solution over the surface. The material is then rinsed in alcohol (isopropyl) for 5-20 seconds to deactivate the active ingredients in the etching solution. Next is a rinse in hot non-chlorinated water for 15-30 seconds, followed by a rinse in hot (70°C) mildly acidic (2-5% acetic acid) water for one minute. The material is then dried with either forced hot air or air dried over night. As a result, the FEP was rendered bondable and therefore able to be sealed with an epoxy.

4.3.3 Epoxy Procedure

A final discovery during the initial testing of the first generation lung Raman probe was the unintended wicking of the epoxy. Medical grade epoxy (Epotek 301, Paisley Products, Ontario CA), was used to seal the probe tip but it traveled up the fibres before curing, resulting in a longer rigid portion at the distal tip. This compounded the problem of advancing through a sharp turn. The wicking was especially difficult to address as the collection fibres are so tightly held together that the inter-fibre space creates a vacuum-like effect for the glue. Reducing the number of fibres would lead to less collection efficiency which is undesirable. Further, reducing the inter-fibre space by having a smaller number of large diameter collection fibres also reduces the flexibility of the catheter, and is also an undesirable option. It was concluded that the glue itself needed to be addressed.
In order to reduce the wicking of the epoxy, the epoxy needed to become more viscous. Therefore it was pre-cured prior to application on the peripheral lung Raman probe. The optimal time chosen for the pre-curing process was done through the development of a ‘dummy’ set-up using two plastic tubes (one inside the other) in order to simulate the small space between the fibre bundle and the FEP covering. Different pre-curing time intervals were tested and the subsequent rigid portion of the tubes were measured after a full curing process. 20 minutes was chosen as the optimal time which allowed for a seal without large wicking potential.

The ability of the epoxy to create a seal with the etched FEP was tested using a phantom set-up. A non-operational fibre bundle (phantom) with an FEP sheath was etched, sealed and tested for leakage. This phantom was first tested by being submerged in water and high pressure air was injected into the proximal end of the phantom. There were no bubbles at the distal end of the phantom, indicating that air was not leaking through the seal. The phantom was then tested by injecting water into the proximal end of the phantom. There was no water leakage at the distal end of the phantom, indicating that the water was not leaking through the seal. The epoxy was therefore pre-cured and applied to the Raman probe tip.
4.3.4 Final Peripheral Probe Design

Through multiple alterations to an original design (Figure 4.1), a functional peripheral lung Raman probe was developed with a diameter of 1.35 mm. The first generation probe underwent ferrule removal, a chemical etching procedure, and a re-gluing procedure before a final polish. The probe was then successfully tested down a bronchoscope to ensure a smooth transition through the sharp turn. The final functional probe design can be seen in Figure 4.4. The final probe tip is shown in Figure 4.5.
Figure 4.4: Final design of the peripheral lung Raman probe.
**Figure 4.5:** End on view (A) and side view (B) of the final functional peripheral lung Raman probe.

### 4.4 Measurements on Tylenol and *in vivo* Skin

Following a successful test down a bronchoscope, the developed Raman probe was tested on a standard Tylenol sample as well as an *in vivo* skin sample to determine if Raman spectra were able to be acquired. The probe was attached to the previously described Raman system (5) (Chapter 2 and 3) and a sample of N-acetyl-para-aminophenol (Tylenol) standard with a fixed probe to sample geometry was measured 100 times. After dark background removal, intensity calibration, smoothing with a boxcar fit, and fluorescence removal (5th order polynomial via VA), the averaged spectra was normalized to the AUC. Figure 4.6 shows the mean Raman spectra of Tylenol collected with the developed peripheral probe.
Figure 4.6: The mean spectra (n=100) of Tylenol collected with the newly developed peripheral lung Raman probe.

Although the Tylenol standard from Chapter 2 was measured in the low frequency (fingerprint) region, there is an overlapping region of 400 cm\(^{-1}\) with the current system. This overlapping region is where the main Raman peaks are found in a Tylenol sample and is thus an appropriate region to compare peak. Figure 4.7 shows the comparison between the central lung probe (Chapter 2) and the newly developed peripheral lung probe.
Figure 4.7: The mean Raman spectra (n=100) of Tylenol collected with the central lung probe (bottom) and the peripheral lung probe (top). Spectra are separated on the y-axis for clarity and only the overlapping region is shown on the x-axis.

It can be seen that the Raman peaks found in Tylenol using the central lung probe were also detected using the peripheral lung probe. The peaks are located at the appropriate wavenumbers and show the same shape as the previous collection. It was concluded that the peripheral Raman probe was functioning similarly to previous systems. The probe was then tested on in vivo location of palm skin of a 28 year old Caucasian volunteer. The spectral peaks
were analysed and assigned appropriate vibrations. Figure 4.8 shows the mean (n=100) of the collected spectra after preprocessing and fluorescence removal. Clear peaks at 1445 cm\(^{-1}\) and 1650 cm\(^{-1}\) have been previously described in the skin (62, 86). These peaks are assigned to the bending mode of CH\(_2\) or CH\(_3\) (proteins and lipids) and the stretching mode of C=O (amide I), respectively. The mean Raman peaks in the high frequency region between 2775 cm\(^{-1}\) and 3000 cm\(^{-1}\) have also been previously described (Chapter 3) (5). The main peak at 2935 cm\(^{-1}\) is assigned to a mixture of CH vibrations in proteins and CH\(_3\) asymmetric stretching modes of lipids and nucleic acids (76, 85).

**Figure 4.8:** The mean (n=100) Raman spectrum of *in vivo* palm skin of a 28 year old Caucasian volunteer measured using the newly developed peripheral probe.
4.5 Chapter Conclusions

A miniature catheter fibre bundle probe with a distal diameter of 1.35 mm capable of navigating the lung architecture to the periphery that has good spectral signal collection efficiency has been developed. The novel probe was created through adaptations and alterations of a first generation design. Following successful tests down the bronchoscope and through the sharp turn, the probe’s optical capabilities were tested on a standard Tylenol sample and \textit{in vivo} palm skin. The probe showed comparable collection efficiency to the previously described central airway probe. These results are promising for the use of Raman spectroscopy as an adjunct diagnostic tool in the peripheral airways. As few technologies are capable of reaching the periphery, novel tools such as this are exciting opportunities to explore new diagnostic strategies. The proof of principle study, showing the first ever collection of peripheral lung Raman spectra, is presented in the next chapter.
Chapter 5: Proof of principle in vivo testing of a Raman spectroscopy system for the detection of peripheral lung lesions

5.1 Chapter Introduction

There is a lack of diagnostic technology available for the periphery of the airways. Moreover, the technologies that are available possess significant limitations on the information they provide to physicians. As the current information provided is strictly physical or morphological, instruments capable of elucidating biochemical characteristics can fill a clinical need. Furthermore, the ability to determine biochemical characteristics in real-time can have implications on the number of biopsies taken – potentially reducing the risk of adverse events, reducing the time needed from physicians, and streamlining a patient’s experience. For example, if a peripheral lesion is biopsied during an REBUS procedure but the sample was non-diagnostic the patient will need another biopsy procedure such as a CT guided transthoracic lung biopsy – a procedure where a biopsy is retrieved by a core needle inserted through the patient’s chest cavity. However, the procedure carries a higher risk for pneumothorax and bleeding. This could be avoided if the nature of the lesion is known through biochemical characterization.

To address the lack of diagnostic technologies for the peripheral airways, we have developed a novel miniature Raman Probe (Chapter 4), which is capable of navigating far into the lung. This miniature probe was tested on ex vivo Tylenol samples and in vivo palm skin, and was concluded to be functioning properly. Therefore, it is hypothesized that it is possible to obtain Raman spectra of the peripheral lungs. Further, we speculate that these spectra are similar to the central lung, such that they show differences between cancer and normal tissue. Here we present, for the first time to our knowledge, the first in vivo clinical test of a peripheral lung Raman system.
5.2 Patient and Methods

This study was approved by the University of British Columbia – BC Cancer Agency Research Ethics Board (certificate number: H06-00010). Patients who were attending the BC Cancer Agency Vancouver Center for a previously scheduled EBUS procedure were invited to volunteer for this study by the attending physician. Patients must have already consented to an EBUS procedure as part of a standard diagnostic procedure or as part of an approved lung cancer prevention study, before being approached to volunteer. Patients were excluded if they had a cardiac pacemaker or implanted defibrillator device, had a known allergic reaction to Xylocaine, were taking a blood thinner such as warfarin or heparin, or had any medical condition such as acute or chronic respiratory failure, which could jeopardize the safety of the patient during participation in the study. Women who were premenopausal were excluded unless they were surgically sterile or on the birth control pill. The guide sheath used at the BC Cancer Agency Vancouver Centre is the Single Use Guide Sheath Kit (Model K-201 KXK, Olympus Medical Systems, Tokyo, Japan). The inner diameter of the guide sheath is 1.4 mm. A total of two patients were measured for this study.

The home-made Raman system has been previously described in detail in Chapter 2 and briefly in Chapter 3. However, it should be noted that the grating was altered to reach the custom range of 1350 to 3050 cm\(^{-1}\). This custom range was chosen because it includes the 1440 cm\(^{-1}\) and 1645 cm\(^{-1}\) peaks. These peaks were not within range for the central lung study (Chapter 3) but were found to have importance in the initial pilot study in 2011 (76). We believe that these additional Raman peaks would help during classification and discrimination as they provide more information for algorithms to train on. Further, during the central lung study we found that there is little information beyond the 3100 cm\(^{-1}\) location so altering the frequency
window would maximize the SNR of the spectra. The maximum excitation power at the tissue surface is 100 mW. The developed miniature Raman catheter, discussed in Chapter 4, consists of one central 100 µm aluminum coated excitation fibre and 31 100 µm surrounding low-OH silica collection fibres. The diameter of the Raman probe is 1.35 mm with a length of 1.7 m. At the proximal end, the excitation channel and collection channels are separated and encased in stainless steel jackets. The distal end of the probe is encased in an FEP tubing.

The Raman clinical procedure is outlined in Figure 5.1. Alteration to standard clinical procedure is minimal with the addition of Raman spectroscopy. The EBUS procedure itself starts with a physician locating a peripheral lung nodule using the radial ultrasound probe encased in a guide sheath. The CT scans provide information on the general location of the lesion. Once a lesion is found, as shown in Figure 5.2 and 5.3, the guide sheath is locked in place to direct further instruments to the proper location. Here, the Raman catheter is inserted into the guide sheath and advanced to the target tissue site. While at the target, multiple (10-15) Raman measurements - to account for tumor heterogeneity - are taken with a 1 second integration time. Once complete, the Raman catheter is removed and the EBUS procedure continues – often by retrieving biopsies from this location. Real time spectral pre-processing included the subtraction of the CCD dark count and application of an intensity calibration factor. The spectra were smoothed by a Savistky-Golay quadratic (82) and fluorescence removal with a fitted polynomial. The pure Raman spectra were then normalized to the area under the curve (AUC).
**Figure 5.1:** The standard clinical procedure for a peripheral lesion needing a biopsy (left) and the Raman procedure (right). Alterations to current clinical workflow are minimal.
Figure 5.2: The clinical procedure for locating and measuring peripheral lesions with the miniature Raman catheter. The top panel shows the localization of a lung nodule with the R-EBUS probe encompassed with the guide sheath. The R-EBUS probe is then retracted while the guide sheath is locked in place. The Raman probe is inserted in the guide sheath and 15-20 measurements are taken. Once the Raman probe is retracted, biopsy forceps are used to retrieve a tissue sample.
Figure 5.3: Size (24.2 mm by 22.9 mm) and location (RB1) of a cancerous peripheral nodule found using EBUS. Point laser Raman spectroscopy measurements were subsequently conducted on this lesion.
5.3 Results and Discussion

Figure 5.4 shows the average peripheral spectra for each pathology group. Both patients had normal tissue measured (n = 31) and one patient had cancerous tissue measured (n=19). Part of the low frequency fingerprint region (1350 – 1800 cm\(^{-1}\)) was obtained during measurements due to the alteration of the detector grating. This region contains Raman peaks in both pathologies. Similarly, the main peaks in the high frequency region (2800 – 3050 cm\(^{-1}\)) are also visible. Differences between the cancerous and normal tissue can be seen in both regions.

The most striking differences can be seen at the Raman peak located at 2887 cm\(^{-1}\), a peak due to the anti-symmetric CH\(_2\) vibrations of lipids (85). I have shown in our prior central lung in vivo study (Chapter 3) that as pathology progresses to invasive lung cancer, the intensity of high frequency lipid peaks weaken (5), similar to our current findings. Furthermore, we have also recently shown for in vivo colon measurements that this peak is reduced as pathology becomes more advanced (123). Agreement between these three different data sets allows us to conclude that the difference seen is due to biochemical changes between lesions. The fact that we are able to obtain clear differences in the Raman spectra in the high frequency region, which are in agreement with previous work, is promising. Peaks at 2930 cm\(^{-1}\) and 2954 cm\(^{-1}\) have been assigned to the CH\(_3\) vibration found in proteins (85).

The peaks in the low frequency region (1350 – 1800 cm\(^{-1}\)) also show differences between the pathology groups. Large peaks are seen at 1550 cm\(^{-1}\) and 1615 cm\(^{-1}\) in the cancerous tissue but these peaks are lessened in the normal tissue. The 1550 cm\(^{-1}\) peak has been assigned as the amide II band, and the 1615 cm\(^{-1}\) peak has been assigned as the C=C stretching mode found in proteins (85). Previous ex vivo murine work by our lab has shown these peaks are seen with a large contribution in blood spectra (124). It is known that increased angiogenesis during tumor
formation is a cancer hallmark (125), however, due to our small sample size we are unable to attribute this change to pathological differences between the tissue. As well, the REBUS probe itself may have caused bleeding during localization – which could have created a blood artefact in the measurement overwhelming the tissue signature. We can only conclude that blood contributes to the cancer spectra more than the normal spectra in the high frequency region – but we are unable to speculate whether that blood signature comes from contained blood inside the tissue or from a bleeding effect due to the procedure. However, differences between the normal and the cancerous tissue can be seen at the 1659 cm\(^{-1}\) peak which was determined to be the amide 1 band (85); which has little signal in whole blood (124). As tissue pathology goes from a normal to a cancerous lesion, this peak becomes less intense, suggesting that biochemical changes are occurring.
Figure 5.4: Mean normalized spectra of each pathology type. Normal (n = 31), Cancer (n=19) and Whole Blood (n=23). Spectra are separated on the y-axis for clarity.

However, whether Raman spectra in this low frequency region are able to separate cancer from benign and normal conditions must be tested with a larger clinical study. The 1442 cm\(^{-1}\) peak was assigned to the CH\(_2\) stretch (85). Similar to our other reported work, no Raman peaks were found between 1800 and 2800 cm\(^{-1}\) (5, 76).

Mann-Whitney U (MWU) statistics were calculated to determine if the normal and cancerous spectra were in fact significantly different. The MWU statistics are plotted in Figure 5.5 superimposed on the difference spectra for two subtractions: Normal – Cancer and Whole Blood – Cancer. Locations with open symbols show wavenumbers where there are significant
differences between the two groups (p-values ≤ 0.005). In the high frequency region (2800 – 3050 cm$^{-1}$) there are significant changes in both data sets around the 2887 cm$^{-1}$ lipid peak, which is in agreement to our earlier studies (5, 123), suggesting that lipid changes occur during pathology progression. In the low frequency region (1350 – 1800 cm$^{-1}$) there are significant changes between 1600 and 1700 cm$^{-1}$ in the Whole Blood – Cancer data set, suggesting that differences due to the presence of the 1659 cm$^{-1}$ amide I band are detectable. Further, there are significant changes in both data sets at the 1550 cm$^{-1}$ and 1615 cm$^{-1}$ bands, suggesting protein differences between pathologies that are not attributed to the increased presence of blood in cancer tissue.
**Figure 5.5:** Mean differences in Raman spectra from peripheral tissue with regions of significant difference (p-value ≤ 0.005) superimposed. Spectra are separated on the y-axis for clarity.
5.4. Chapter Conclusion

We have acquired for the first time, in vivo Raman spectra of peripheral lung lesions which show differences between normal and cancerous tissue. Through development of a novel miniature fibre bundle probe, clear Raman peaks were obtained with good efficiency in the high frequency region as well as part of the low frequency region. Though these results are preliminary and more clinical data is required to confirm these findings, we have shown the feasibility of using Raman spectroscopy as an adjunct device during R-EBUS procedures.

The collected spectra show promise in classifying different pathologies and open up the possibility of using Raman to provide more information in real time to clinicians and potentially reduce the number of biopsies needed during peripheral lung procedures. Further, there remains the possibility of developing a side-firing Raman probe for the peripheral airways to more accurately match the configuration of the REBUS probe. Though this development is technically challenging, we have shown that the use of Raman in the lung periphery is possible. Therefore, probe and component optimization should be considered in the next step.
Chapter 6: Case study: *in vivo* Raman spectrum of an Aspergilloma

6.1 Chapter Introduction

During the course of data collection for the proof of principle *in vivo* peripheral study (Chapter 5), an *in vivo* measurement was acquired from a patient with an Aspergilloma. This is a fungal infection by the species *Aspergillus fumigatus* that presents after a spore is inhaled and logged into a pre-existing lung cavity (126). Active immune systems will engulf this spore before it grows into hyphae, but in lungs that have a suppressed immune system this does not happen (126). The spore will then grow into a ball of hyphae which is visible on LDCT scans (127).

Though this Aspergilloma is a very specific example, lung cancer is not the only disease visible on a LDCT scan. Non-invasive diagnostic techniques like Raman can provide more information to clinicians in real time. The presence of scar tissue, infections and inflammation is often confounding during an endo-bronchial biopsy procedure. These conditions require different treatments than lung cancer, and in some cases like a fungal infection, waiting multiple days for a cytology or pathology report can be too long for a patient who is drastically ill. The ability to classify peripheral nodules into normal, benign or malignant categories could help clinicians narrow the scope of treatment, or start treatment immediately.

However, to create such algorithms, many clinical measurements need to be done. In this case, I had access to a single patient who presented with the condition. Although little information can be obtained with only a single measurement, what can be determined is the feasibility of using Raman Spectroscopy to locate benign *in vivo* conditions. Analyzing this *in*
vivo spectrum to determine the fungal and tissue bio-chemicals contributing to the Raman peaks can help decide if the Raman spectra of infections are unique to the infecting agent. If the Raman peak in vivo contained the same peaks as a pure in vitro sample of Aspergillus fumigatus then follow up with a larger sample set and more clinical data is warranted.

However, there is little information on the Raman spectra of Aspergillus fumigatus. One study has tested if Raman spectroscopy measurements were sensitive enough to measure differences in A. fumigatus and A. lentulus strains. Although the study concluded this was not possible, a spectra was not shown in the publication and spectral peaks were not assigned (128). Other research has looked into using Raman on in vitro samples of other fungi, but the collection has historically been in the low frequency region or on spores which are often not found in Aspergillomas (129-131). Therefore, in order to find a reference sample of the fungus our lab reached out to Dr. Margo Moore in the department of microbiology at Simon Fraser University, who was able to provide us with 5 samples of Aspergillus fumigatus which could be subjected to bench-top optical testing and subsequently compared to the collected in vivo measurement.

6.2 Materials and Methods

6.2.1 In vivo Measurement

The in vivo Aspergilloma measurement was collected using the previously described Raman system and miniature peripheral catheter (Chapter 5). 13 measurements were taken in vivo and preprocessing of the spectra included a dark background removal, application of an intensity calibration, fluorescence background removal with the Vancouver Algorithm (80), computing of the mean spectra (n=13), and normalizing with the AUC.
6.2.2 In vitro Fungal Measurements

Five samples of *Aspergillus fumigatus* were grown on five different media. The samples were grown in the biology department at Simon Fraser University in the lab of Dr. Margo Moore. The five types of media used were 1) RPMI with goat serum, 2) YAG, 3) Neiland’s media with iron, 4) Neiland’s media without iron and 5) RPMI without goat serum. These media were chosen in order to represent different environments (nutrients) and provide a variety of samples. Roswell Park Memorial Institute Medium (RPMI) has a large amount of phosphate, Yeast Agar Glucose (YAG) contains carbohydrates (glucose) and Neiland’s Media provides (or removes) sources of Iron for the fungus. The samples were grown simulating *in vivo* conditions as much as possible, meaning the fungus was grown into hyphae balls. It is rare for *A. fumigatus* to produce spores when it is inside the lung, so the samples were grown in a warm environment to suppress sporulation (which is a survival mechanism in harsh environments). However, after inspection of the samples, sample 5 (RPMI without serum) in fact did grow spores and was therefore excluded from the experiments. The media which closely resembles the conditions of the peripheral airways is the RPMI with goat serum.

Once the volume of hyphae reached an adequate level for optical measurements (5 ml) the hyphae was washed with distilled water (to remove excess media) and flash frozen (sample one, RPMI with goat serum, remained fresh). The samples were then loaded into specially designed airtight optical containers. The mechanical drawing of the designed optical containers is shown in Figure 6.1 and Figure 6.2. The loaded containers are shown in Figure 6.3 A-D. The optical containers consisted of a 3D printed base that locked with a cap after the samples were placed inside. A thin quartz optical window was sealed to the cap using Epotek medical epoxy
(Epotek 301, Paisley Products, Ontario CA). Raman spectroscopy measurements were taken through the thin (0.2 mm) quartz window (Electron Microscopy Sciences, Hatfield, PA). The samples were transported to the BC Cancer Research Centre for testing.

The four viable samples were subsequently measured using the peripheral lung Raman system (PLRS) which is described in detail in Chapter 5. 100 measurements of each sample were taken using the PLRS. Following the fungal measurements, each media type was also measured using the PLRS. Measurements were background corrected, intensity calibrated, fluorescence removed (Vancouver Algorithm), and averaged.

Figure 6.1: The mechanical design of the airtight optical container base used to hold the *A. fumigatus* samples. This base locked with the cap shown in Figure 6.2.
Figure 6.2: The mechanical design of the airtight optical container cap used to hold the *A. fumigatus* samples. This cap locked with the base shown in Figure 6.1. The window (diameter 11.50 mm) was covered and sealed with a thin quartz window.
Figure 6.3: *Aspergillus fumigatus* samples loaded into specially designed airtight optical containers. The samples were grown on (A) RPMI with goat serum, (B) YAG, (C) Neiland’s Media with Iron and (D) Neiland’s Media without Iron.
6.3 Results and Discussion

6.3.1 *In Vivo* Measurement with the Peripheral Lung Raman System

Figure 6.4 shows the normalized mean Raman spectra obtained *in vivo* from an Aspergilloma. The spectrum shows clear Raman peaks in both the low (1500-1800 cm\(^{-1}\)) and high frequency (1800-3000 cm\(^{-1}\)) regions. The peak at 1598 cm\(^{-1}\) was assigned to the C=O stretching vibrations of peptide (protein) backbones (phenylalanine and tyrosine) (85). The peak at 1726 cm\(^{-1}\) was assigned to the C=O stretching vibrations of amino acids and lipids (85). Peaks in the high frequency region were found at 2887 cm\(^{-1}\) which is assigned to the anti-symmetric CH\(_2\) vibrations of lipids, as well as 2930 cm\(^{-1}\) and 2954 cm\(^{-1}\) which are the CH\(_3\) vibration found in proteins (85).
Figure 6.4: The *in vivo* mean (n=13) spectra of an aspergilloma. The peaks of interest at 1598 cm\(^{-1}\) and 1726 cm\(^{-1}\) are shown in dashed vertical lines.

The high frequency region peak assignments are described in detail in Chapter 3. There are two peaks of interest in the low frequency region (at 1598 cm\(^{-1}\) and 1726 cm\(^{-1}\)) that were not present in the normal *in vivo* tissue measurements from the peripheral airways, which are described in detail in Chapter 5. These peaks have been assigned to amide I stretches in biological tissue (85), but at the time of writing, there are currently no *in vivo* fungal measurements in the high frequency region for us to compare to.
6.3.2 *In Vitro* Measurements with the Peripheral Lung Raman System

Figure 6.5 shows the mean Raman spectra of each fungal sample acquired with the PLRS. Clear Raman peaks are seen in the low frequency (1500-1800 cm\(^{-1}\)) and high frequency regions (1800-3000 cm\(^{-1}\)). The peak at 1598 cm\(^{-1}\) is in the amide I band of stretching vibrations mostly due to C=O (85), found in phenylalanine, tyrosine and cytosine. The peak at 1726 cm\(^{-1}\) is as well assigned to the C=O stretching vibration of amino acids and lipids. The high frequency band (2700 cm\(^{-1}\) to 3100 cm\(^{-1}\)) is due to CH, CH\(_2\) and CH\(_3\) stretches found in lipids and proteins. This region is described in detail in Chapter 3.

Peaks of note include the 1598 cm\(^{-1}\) peak and the 1726 cm\(^{-1}\) peak, as these were also located in the *in vivo* Aspergilloma measurement. Further, there is relatively little fluctuation between the four *A. fumigatus* samples, although sample 1 (RPMI with goat serum) has less intensity than the others. This sample is the only fresh sample, but the relative shape of the spectra remains stable across all the samples. The consistency among the four sample types is expected as they are the same species and strain (differing only in the media they grew on).

Each of the four media types were also measured using the PLRS. The mean spectra for each media type can be seen in Figure 6.6. Although there are peaks present in the media, they are at different locations than the ones identified in the fungal samples. This is expected as the samples were washed in distilled water to remove excess media prior to freezing. It was concluded that the media was not contributing to the *in vitro* spectra of the fungal samples.
**Figure 6.5:** The mean (n=100) *in vitro* spectra of the four *Aspergillus fumigatus* samples acquired with the PLRS. The peaks of interest at 1598 cm⁻¹ and 1726 cm⁻¹ are shown in dashed vertical lines.
Figure 6.6: Mean spectra using the PLRS for each media type: 1) RPMI with serum, 2) YAG, 3) Neiland’s with Iron and 4) Neiland’s without Iron. Dashed lines showing peaks of interest at 1507 cm\(^{-1}\), 1686 cm\(^{-1}\), 1800 cm\(^{-1}\) and 1950 cm\(^{-1}\).

6.3.4 Spectral Comparison

The similar region of the three mean spectra (in vivo Aspergilloma, in vivo normal, and in vitro fungal sample) is shown in Figure 6.7. The in vitro fungal sample spectrum shows the overall mean, so the 100 measurements taken with the PLRS per sample has been averaged together to produce a mean of 400 samples. There are striking differences in the amount of noise found in the in vivo Aspergilloma sample, which is mostly likely due to the low efficiency of the
detector in the long wavelength region. This low efficiency is offset by acquiring many measurements (n=400) during spectral acquisition of the *in vitro* fungal samples, however during *in vivo* collection 400 spectra are not possible as this would take too much time during a routine procedure. Nevertheless, the two peaks at 1598 cm\(^{-1}\) and 1726 cm\(^{-1}\) found in the fungal samples line up well with the peaks found *in vivo*.

**Figure 6.7:** The mean spectra of the *in vivo* Aspergilloma measurement (n=13), the *in vivo* normal lung tissue measurement (n=15) and the *in vitro* fungal sample measurement using the PLRS (n=400).
The three spectra show some similarity in the Raman peaks. Specifically it can be seen that the peak at 1726 cm\(^{-1}\) is present in the Aspergilloma spectrum and the fungal samples spectrum but not the normal lung tissue spectrum. Though much more data needs to be collected before making any statistical inference, this small data set suggests that there are real measurable differences in the Raman between an Aspergilloma and normal lung tissue based on true fungal signatures. This 1726 cm\(^{-1}\) peak appears to be promising for the discrimination of Aspergilloma tissue in the lung periphery.

6.5 Chapter Conclusion

This case study is limited by the single in vivo measurement obtained, but in vitro measurements taken from Aspergillus fumigatus samples show good agreement. Though it is probable that other important factors, like immunology changes due to a fungal infection, would influence any interpretation, it can be seen that some in vitro fungal peaks align with the peaks seen in vivo. The acquired in vivo Aspergilloma spectrum is complex, yet, it remains that peaks seen at 1598 cm\(^{-1}\) and 1726 cm\(^{-1}\) in vitro are also seen in vivo, and a follow up study is warranted.
Chapter 7: Conclusions and Future Directions

7.1 Conclusions

This dissertation addressed the hypothesis that real time rapid Raman Spectroscopy can be introduced into clinical procedures as an adjunct method capable of detecting the biochemical differences among normal and various abnormal tissue, which can then be used to classify tissue in real time and detect premalignant and malignant lesions in the lung. I sought to address three specific aims: to 1) validate the performance of RS for in vivo pre-malignant and malignant classification in the central airways, 2) design a novel miniature Raman probe for use in the peripheral airways and 3) determine the feasibility of using RS for in vivo lesion classification in the peripheral airways.

Following hardware optimization and development of analytical techniques, the performance of RS for early cancer detection in the central airways was validated through a large single centred clinical study. This study compared the performance of Raman spectroscopy for classifying lung lesions into benign (Mild Dysplasia and better) and premalignant and malignant (Moderate Dysplasia and worse) categories. Measurements of over 80 patients yielded 280 spectra that were subsequently classified using multivariate mathematical models. If feature selection methods (Forward Stepwise, LASSO, GA) and dimension reduction methods (PCA or PLS) are used prior to developing a regression model RS is able to detect malignant lung lesions with a high sensitivity (90%) and good specificity (65%). Compared to the visual assessment of lesions alone, RS is able to improve the sensitivity of localizing HGD and malignant lesions by 1.89 times when compared to WLB+AFB alone (visual grade 3 based diagnoses). The inclusion of RS also improved the specificity of localizing HGD and malignant lesions by 3.8 times (visual
grade 2 and 3 combined diagnosis). The removal of subjective interpretation by clinicians remains one of the largest foreseeable benefits to using RS in the central airways.

These results also suggest that the inclusion of RS to routine clinical procedures in the lung can provide benefit to the patient and should be seriously considered as an adjunct device. Presumably the same would be true in the peripheral airways. However, there are significant differences in the structure of the periphery when compared to the central airways, most importantly the smaller diameter, and the sharp turns to get into the upper lungs. This warranted a novel miniature lung probe designed specifically for the peripheral airways. A re-useable fibre bundle probe with a diameter of 1.35 mm was created. Though the preliminary design was based on the central lung probe, initial testing of the probe through the instrument channel of a bronchoscope yielded negative results (the probe was unable to make the journey through the bronchoscope). As a result, multiple design changes were implemented, such as the removal of a metal ferrule at the distal tip. The fabrication procedure was also altered to include the chemical etching of the FEP and the pre-curing of the medical grade epoxy before sealing the distal tip. These changes then yielded positive results as the probe successfully made the journey through the bronchoscope. Subsequently, the new novel miniature Raman probe was tested optically on a Tylenol standard and in vivo palm skin. Through comparisons between previous probe designs and the novel probe it was concluded that the novel probe was producing acceptable Raman spectra.

Finally, this novel miniature Raman probe was tested in vivo to determine the feasibility of using RS for the classification of peripheral lung nodules. The first in vivo Raman spectra of peripheral airways were acquired from two different patients. Clear Raman peaks can be obtained from both cancer and normal peripheral lung tissue. Significant differences in
peripheral cancer and peripheral normal tissue are found when using a MWU test. Further, a single case of peripheral Aspergilloma was acquired *in vivo*. The Raman spectra of this case showed similarities to *in vitro* fungal samples and presents the opportunity to classify benign nodules found in the lung periphery. Though many more cases are needed to develop clinical conclusions about using RS for lesion classification in the peripheral airways, I have shown that clear Raman peaks can be obtained from a variety of conditions and it is feasible to follow up with a pilot clinical study.

### 7.2 Current Limitations and Future Directions

The data collected, analysed and presented from the large central airway study (Chapter 3) is from a single health care centre. Though I collected a well diverse set of patients, lesions and final pathologies, there remains selection bias in the data. Each patient was previously scheduled for a bronchoscopy at the BC Cancer Agency Vancouver Centre, meaning a radiologist or family clinician previously inspected the scan and then suspected the diagnosis of lung cancer. This suggests that the lesions in the study were more likely to be diagnosed as cancer than lesions found in non-specialized hospitals. Further, the patient was then sent to a lung cancer specialist who had significant expertise in bronchoscopy procedures, which may not represent all respirologists. However, the BCCA Vancouver Centre is located on the west coast of British Columbia, Canada, where there is significant diversity among the population. Large population subsets of Asian, Middle Eastern and European peoples call Vancouver home. Although these specific demographics of the clinical study have not been analyzed, it is probable that the data set represents ethnicities from all over the world. This would mean that the
conclusions and outcomes of the clinical study may show realistic and transferable results that can be extrapolated worldwide.

To combat the aforementioned limitations and to confirm the results of the clinical study using RS in the central airways, a large multi-centred clinical trial should be conducted. This should include a variety of care sites (specialist and non-specialist hospitals), respirologists with varying levels of expertise, and different locations around the country or the globe.

The data collected and presented from the initial *in vivo* testing of a miniature peripheral Raman probe (Chapter 5 and 6) was very preliminary. A total of three patients were measured (2 in Chapter 5 and one in Chapter 6), and one should be very cautious and conservative when drawing conclusions based on such a small sample size. Though significant differences are seen between cancer and normal tissue, this finding must be validated further with more measurements. A pilot study similar to the one conducted in the central airways by Short et al. in 2011 (76) would be an appropriate direction forward and would help elucidate the differences in peripheral nodules. Though, I strongly believe that RS has great potential as an adjunct device in the peripheral airways, clinical data can help make that a reality.
Bibliography


10.1097/LBR.0b013e3181da2ca8.


The Link Between Inflammation and Cancer

Improved Group Type Analysis, Appl Spectrosc (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy, Bioinformatics, 23, 2507-17.


Huang, N., Short, M., Zhao, J. et al. (2011) Full range characterization of the Raman spectra of organs in a murine model, Optics express, 19, 22892-22909.


## Appendices

### Appendix A

<table>
<thead>
<tr>
<th>Stage</th>
<th>TNM Grading</th>
<th>Description</th>
</tr>
</thead>
</table>
| 0     | Tis N0 M0   | Lesions are pre-invasive lesions known as CIS with no lymph node involvement  
Do not have distant metastases |
| IA    | T1a N0 M0   | Tumors is 3cm or less with no lymph node involvement  
Do not have distant metastases |
|       | T1b N0 M0   | Tumors is 3cm or less with no lymph node involvement  
Do not have distant metastases |
| IB    | T2a N0 M0   | Tumor is between 3cm and 5cm with no lymph node involvement  
Do not have distant metastases |
|       | T2b N0 M0   | Tumor is between 3cm and 5cm with no lymph node involvement  
Do not have distant metastases |
| IIA   | T1a N1 M0   | Tumor is 5cm or less with metastatic cells found in the ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes  
Do not have distant metastases |
|       | T1b N1 M0   | Tumor is 5cm or less with metastatic cells found in the ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes  
Do not have distant metastases |
|       | T2a N1 M0   | Tumor is 5cm or less with metastatic cells found in the ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes  
Do not have distant metastases |
| IIB   | T2b N1 M0   | Tumor is between 5cm and 7cm with metastatic cells found in the ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes  
Do not have distant metastases |
|       | T3 N0 M0    | Tumor is greater than 7cm with no lymph node involvement  
Do not have distant metastases |
| IIIA  | T1a N2 M0   | Tumor is 7cm or less with metastatic cells found in the ipsilateral mediastinal and/or subcarinal lymph nodes of the mediastinum (central area of the chest) on the same side as the primary tumor  
Do not have distant metastases |
|       | T1b N2 M0   | Tumor is 7cm or less with metastatic cells found in the ipsilateral mediastinal and/or subcarinal lymph nodes of the mediastinum (central area of the chest) on the same side as the primary tumor  
Do not have distant metastases |
|       | T2a N2 M0   | Tumor is 7cm or less with metastatic cells found in the ipsilateral mediastinal and/or subcarinal lymph nodes of the mediastinum (central area of the chest) on the same side as the primary tumor  
Do not have distant metastases |
|       | T2b N2 M0   | Tumor is 7cm or less with metastatic cells found in the ipsilateral mediastinal and/or subcarinal lymph nodes of the mediastinum (central area of the chest) on the same side as the primary tumor  
Do not have distant metastases |
T3 N1 M0  
T3 N2 M0  

Tumor is greater than 7cm with metastatic cells found in the ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes OR

Tumor is greater than 7cm with metastatic cells found in the ipsilateral mediastinal and/or subcarinal lymph nodes of the mediastinum (central area of the chest) on the same side as the primary tumor

Do not have distant metastases

T4 N0 M0  
T4 N1 M0  

Tumor is any size, has invaded major body organs with no lymph node involvement OR

Tumor is any size, has invaded major body organs with metastatic cells found in the ipsilateral peribronchial and/or ipsilateral hilar lymph nodes and intrapulmonary nodes

Do not have distant metastases

T1a N3 M0  
T1b N3 M0  
T2a N3 M0  
T2b N3 M0  

Tumor is 7cm or less with metastatic cells found in the contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph nodes

Do not have distant metastases

T3 N3 M0  

IIIB  

Tumor is greater than 7cm with metastatic cells found in the contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph nodes

Do not have distant metastases

T4 N2 M0  
T4 N3 M0  

Tumor is any size, has invaded major body organs with metastatic cells found in the contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph nodes

Do not have distant metastases

IV  
Any T Any N M1a or M1b  

Do have distant metastases
<table>
<thead>
<tr>
<th>Generation</th>
<th>Location</th>
<th>Diameter (mm)</th>
<th>Cartilage</th>
<th>Tissue Layers</th>
<th>Epithelium</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Trachea</td>
<td>18</td>
<td>U-Shaped Rings stacked consecutively and connected by smooth muscle</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Pseudo-stratified columnar ciliated including three cell types – basal, ciliated and goblet cells</td>
</tr>
<tr>
<td>1</td>
<td>Right and Left Main Bronchi leading to each lung</td>
<td>12</td>
<td>Irregular bands</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Pseudo-stratified columnar ciliated including three cell types – basal, ciliated and goblet cells</td>
</tr>
<tr>
<td>2 to 3</td>
<td>Lobar Bronchi leading to upper, middle*, and lower lobes</td>
<td>8 to 5</td>
<td>Irregular bands</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Pseudo-stratified columnar ciliated including three cell types – basal, ciliated and goblet cells</td>
</tr>
<tr>
<td>4</td>
<td>Segmental Bronchi</td>
<td>4</td>
<td>None</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Pseudo-stratified columnar ciliated including three cell types – basal, ciliated and goblet cells</td>
</tr>
<tr>
<td>5 to 11</td>
<td>Small Bronchi</td>
<td>3 to 2</td>
<td>None</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Pseudo-stratified columnar ciliated including three cell types – basal, ciliated and goblet cells</td>
</tr>
<tr>
<td>12 to 14</td>
<td>Small Bronchi, Bronchioles and Terminal bronchioles</td>
<td>2 to 0.7</td>
<td>None</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Cuboidal</td>
</tr>
<tr>
<td>15 to 18</td>
<td>Respiratory Bronchioles</td>
<td>0.4</td>
<td>None</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Cuboidal transitioning to flat cells between alveoli</td>
</tr>
<tr>
<td>19 to 22</td>
<td>Alveolar Ducts</td>
<td>0.3</td>
<td>None</td>
<td>1) mucosal epithelium 2) submucosa (gland layer) 3) cartilage and muscle layer</td>
<td>Simple squamous epithelium (type I and type II pneumocytes)</td>
</tr>
<tr>
<td>23</td>
<td>Alveoli</td>
<td>0.2</td>
<td>None</td>
<td>1) single celled mucosal layer</td>
<td>None</td>
</tr>
</tbody>
</table>

*Middle lobe is present only in the Right Lung
**Table A.3: Pathology Details for Premalignant Squamous Cell Central Airway Lesions (27, 133, 134).**

<table>
<thead>
<tr>
<th>Mucosal Lesion</th>
<th>Progression of Mature Cells</th>
<th>Basal Layer</th>
<th>Luminal Surface</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperplasia (Goblet cell)</td>
<td>All cells achieve maturation</td>
<td>Normal</td>
<td>Goblet cell proliferation such that goblet cells are adjacent to each other</td>
<td>Generally thought to be reversible</td>
</tr>
<tr>
<td>Hyperplasia (Basal cell)</td>
<td>The cells continue to mature when progressing to the luminal surface</td>
<td>Basal cell proliferation</td>
<td>Goblet and cilia cells remain normal at the luminal surface</td>
<td></td>
</tr>
<tr>
<td>Metaplasia</td>
<td>Continuous maturation of cells from the basal layer to the luminal layer.</td>
<td>Transformation of columnar epithelium to squamous type</td>
<td>Keratinized epithelial cells at the luminal surface</td>
<td></td>
</tr>
<tr>
<td>Mild Dysplasia</td>
<td>Only partial maturation of cells as they progress to the luminal surface.</td>
<td>The basal layer is expanded to 1/3 of the total mucosal layer due to cellular crowding</td>
<td>Cells at the luminal layer experience flattening</td>
<td>Majority will regress</td>
</tr>
<tr>
<td>Moderate Dysplasia</td>
<td>Little maturation of cells as they progress to the luminal surface.</td>
<td>The basal layer is expanded to 2/3 of the total mucosal layer due to cellular crowding</td>
<td></td>
<td>Higher probability of Invasive Cancer development somewhere in the lung</td>
</tr>
<tr>
<td>Severe Dysplasia</td>
<td>No maturation of cells as they progress to the luminal surface.</td>
<td>The basal layer is expanded to the vast majority of the total mucosal layer due to cellular crowding</td>
<td></td>
<td>Higher probability of progressing to an invasive lesion</td>
</tr>
<tr>
<td>Carcinoma in situ</td>
<td></td>
<td>Basement membrane has been breached</td>
<td>Squamous patterns present - intercellular bridging and/or keratinisation of individual cells</td>
<td>Invasive lesion</td>
</tr>
<tr>
<td>Invasive Carcinoma</td>
<td></td>
<td>Basement membrane has been breached</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

A detailed example of the SG smoothing algorithm and calculation of the second order derivatives can be seen below.

**Figure B.1:** An example raw spectrum collected from *ex vivo* Rat Liver tissue.
The collected raw spectrum is first calibrated for the spectral response of the instrument using the known spectral irradiance of a standard tungsten lamp (RS-10A; Gamma Scientific, San Diego, CA).

The calibrated raw spectrum is then smoothed using the Savitzky-Golay algorithm (SG). This algorithm performs a least squares fit of the data to a polynomial in a subset (window) of the data. Using SG for a smoothing filter is similar to performing a moving average, which is represented by equation B.1 below which has been adapted from (135). \( X_i \) and \( X_i' \) are data points (in our case spectral intensity) of the original (raw) spectrum and the smoothed spectrum, respectively. The values \( \omega_j \) are the weighting factors of each data point in the defined window. The defined window is odd (2n+1) in order to define a central value. This central value is then replaced by the calculated average within the entire window.

\[
X_i' = \frac{\sum_{j=-n}^{n}(x_i + j \cdot \omega_j)}{\sum_{j=-n}^{n}(\omega_j)} \quad \text{(B.1)}
\]

In a simple moving average, the weightings of each data point in the window would be 1 (\( \omega_j=1 \)), as a simple moving average treats each of the data points with equivalent weight. Although when using this for spectral smoothing, problems often arise when small Raman peaks are collected. If the peaks are small enough to be overwhelmed by the neighbouring data points, they may be averaged out during the smoothing process, and a loss of information occurs. In 1964 Savitzky and Golay published a modified version similar to the moving average, but SG performs a least squares fit of the data to a polynomial function which is convex (in our case a quadratic). Since this polynomial is convex, there will be weighted factors in the average. The central value of the convex polynomial will hold more weight than the peripheral values. This has shown to preserve valuable spectral peak information as small Raman peaks will not be
completely averaged out. Although one must still choose a window size for the moving SG filter. Smaller window sizes (like 5 or 7 points) may not smooth the spectra sufficiently and noise will still be prominent. However, large window sizes (like 23 or 25) may in fact over smooth the spectra resulting in information loss as Raman peaks may be averaged out. We have found a window size around 11 or 13 works well without compromising the amount of Raman information collected.

\[
   x'_i = \frac{\sum_{j=-n}^{n}(x_{i+j} \omega_j)}{\sum_{j=-n}^{n} \omega_j}
\]  

Equation B.1

The SG filter will therefore add weightings (\(\omega_j\)) to equation B.1. The calculation of these weightings (often called SG coefficients) can be seen in detail in ref. (135) and tables of possible coefficients were published in the original Savitzky-Golay paper (82). An example of the SG filter calculation for 3 different window sizes can be seen in Figure B.2.
Figure B.2: Example calculation showing the SG smoothing function over 3 different window sizes. It should be noted that because the SG is using a window size to estimate the value of the central point, data values at the edges of the spectrum are unable to be estimated and thus the spectra is truncated slightly.

Figure B.3 shows the example raw spectra (in Figure B.1) after application of the SG smoothing filter with a window size of 13 points. The spectrum is showing the region of interest (2100 – 3500 cm\(^{-1}\)) as an example.
Once the spectra are smooth, a second order derivative is applied to remove the background fluorescence. Second order derivatives are valuable methods of background removal when the fluorescence is very high. From above we can see that SG is able to calculate a smoothed spectrum based on a simplified least squares fitting to the original raw data. The SG is also able to calculate the derivative of the spectra. Details of this process can be seen in (136) and the SG coefficients for the second order derivative were also taken from (82). Figure B.4 shows a sample calculation of the second order
derivative of a spectrum using an 11 point window size, and Figure B.5 shows the example spectrum from Fig. B.3 after the application of SG 2\textsuperscript{nd} order derivatives (in the region of interest).

\begin{figure}
\centering
\includegraphics[width=\textwidth]{example_spectrum}
\caption{Example calculation showing the SG 2nd order derivative function over an 11 point window size. It should be noted that because the SG is using a window size to estimate the value of the central point, data values at the edges of the spectrum are unable to be estimated and thus the spectra is truncated.}
\end{figure}
Figure B.5: An example 2\textsuperscript{nd} order derivative spectrum collected from \textit{ex vivo} Rat Liver tissue.

To complete the pre-processing steps, the spectra are then normalized. In this example, they will be normalized by summing the squared derivative values and dividing each variable by that sum. An example of this calculation can be seen in Figure B.6, and the example Rat Liver spectrum in the region of interest can be seen in Figure B.7 after the normalization procedure.
**Figure B.6:** Example calculation showing the normalization procedure of the SG 2nd order derivatives.
Figure B.7: An example of a normalized 2\textsuperscript{nd} order derivative spectrum collected from \textit{ex vivo} Rat Liver tissue.
Appendix C

Waveband Selection via STEP

STEP is a systematic method for constructing a multiple linear regression model. A multiple linear regression model will create a linear mathematical model that attempts to explain the relationship of multiple prediction variables on a single response variable. In our case, the model will try to explain the relationship between spectral intensity difference at various wavenumbers and the histopathology grouping of the tissue. This is done by fitting a linear equation to the data set, such that the residual mean square error is minimized. This linear equation can be seen in equation C.1 and is modified from (137).

\[
\hat{y}_i = b_0 + b_1 x_{i,1} + b_2 x_{i,2} + \ldots + b_{p-1} x_{i,p-1} \quad (C.1)
\]

In which \( \hat{y}_i \) is the estimated y value (histopathology group), \( b_0 \) to \( b_{p-1} \) are regression coefficients (weightings of each term), \( x_1 \) to \( x_{p-1} \) are the terms of the model (wavenumbers), \( p \) is the number of terms in the model and \( i \) is the \( i \)th individual member or unit in the model. The STEP algorithm optimizes this linear model by including or eliminating various terms \( (x_i – \)wavenumbers) which improve the model fit to the data. In other words, the algorithm will create a model in which the Root Mean Square Error (RMSE) is minimized. The equation for the RMSE can be seen in equation C.2.

\[
RMSE = \sqrt{\frac{SSE}{n-1}} \quad (C.2)
\]
In which the numerator is known as the Sum of Squared Errors (SSE) and is defined in equation C.3.

\[ SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \]  

(C.3)

In which \( n \) is the sample size, \( y_i \) is the observed (true) data value, and \( \hat{y}_i \) is the predicted data value by the model. In our case, \( y_i \) is the true histopathology grouping of a spectrum, and \( \hat{y}_i \) is the predicted histopathology grouping calculated by the model. Therefore, the smallest difference between the histopathology predicted by the multi-linear model and the true histopathology will be the model with the smallest RMSE. In other words, the RMSE is the standard deviation of the residuals of the model – the smaller RMSE, the better the model is fit to the data.

This can be used to select wavebands, or wavenumbers that have important contributions to the histopathology groupings. The wavenumbers that have a spectral intensity difference which contribute most to the histopathology grouping will have a high regression coefficient (a high \( b \)– value), where small coefficients indicate that the spectral intensity difference at that wavenumber is not as important for the histopathology group. But one can see that when all predictor variables are entered into the model at once, the computational requirements can be high. Further, determining which regression coefficients are sufficiently low enough to leave out of the model and still have an accurate prediction can be challenging. To combat this, we have used Stepwise multiple regression. STEP is systematic iterative routine that creates a multi-linear model (equation C.1) in which subsequent forms of the model have a lower RMSE (equation...
C.2) than previous versions. Specifically, the algorithm consists of the following three steps (adapted from ref. (138)):

1. Fit the initial model.

2. Determine if the P values (of the F statistic) of any terms not in the model are below the entrance requirement, if so, add the most significant one to the model; otherwise, go to step 3.

3. Determine if the P values (of the F statistic) of any terms in the model are above the exit requirement, if so, remove the least significant one from the model, go to step 2; otherwise, end.

In which commonly used entrance requirement are 0.05 or 0.01 and exit requirement are 0.1 (138). It can be seen that the P values (calculated through hypothesis testing at each term in the model) will be recalculated with every iteration or step of the model. The model is stable when there are no terms which violate the exit or entrance requirements. However, this model (and thus the terms which are included) will change based on the spectra which are used for training (to develop) the model. As discussed in Chapter 3, waveband selection results depend on the sample spectra and sample size. In order to find reliable optimal wavebands, a leave one out cross validation (LOO-CV) protocol was used. In the LOO-CV waveband selection procedure, a single spectrum was left out with the remaining spectra used for waveband selection operation. A set of wavebands were selected which gave the best diagnostic performance of the training spectra. By repeating this procedure, every spectrum was left out once for wavelength selection purpose. At the end, n sets of wavebands were selected, where n was the total number of cases.
The n sets of wavebands were then accumulated. The wavebands with higher odds from the LOO-CV analysis were chosen for subsequent PCA-DA and PLS analyses.

**Dimension Reduction via PCA**

For PC-GDA analysis, we followed the methodology in (99). As discussed in Chapter 3, the training process consisted of the following procedures: (1) the mean and standard deviation of the training spectra data set were calculated, shown in equation C.4 and C.5 respectively, (2) each spectrum in the training data set was standardized by subtracting the mean and then dividing by the standard deviation, (3) the standardized training spectra were analyzed with principal component analysis, discussed in length below. PC factors of the training cases and PC loadings were obtained, and (4) a generalized (linear) discrimination (GDA) model was developed from the PC factors which could be used directly to predict the new cases – also discussed in length below.

\[
\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}
\]  
(C.4)

In which \(\bar{X}\) is the mean of the data set. \(X_i\) is each individual data point, in our case the intensity value at each wavenumber, \(n\) is the number of spectra, and \(i\) represents the ith member (or spectra) in the group.
In which \( s \) represents the standard deviation – the average distance of the individual data points to the mean of the group. In our case, this is the average distance of the spectral intensities at each wavenumber to the mean at the same wavenumber. Once the spectra are standardized with equation C.4 and C.5 in the aforementioned steps (1) and (2), they are analyzed with Principal Component Analysis (PCA). PCA is an orthogonal transform that maps the original data into a lower dimension space where the variance among predictive features is maximized (98). Here, linear combinations that explain most of the variance in the data are extracted by calculating the covariance matrix of the data (equation C.6 - C.9) then calculating the eigenvectors and eigenvalues of that covariance matrix (139).

\[
s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}} \quad \text{(C.5)}
\]

In which \( s \) represents the standard deviation – the average distance of the individual data points to the mean of the group. In our case, this is the average distance of the spectral intensities at each wavenumber to the mean at the same wavenumber. Once the spectra are standardized with equation C.4 and C.5 in the aforementioned steps (1) and (2), they are analyzed with Principal Component Analysis (PCA). PCA is an orthogonal transform that maps the original data into a lower dimension space where the variance among predictive features is maximized (98). Here, linear combinations that explain most of the variance in the data are extracted by calculating the covariance matrix of the data (equation C.6 - C.9) then calculating the eigenvectors and eigenvalues of that covariance matrix (139).

\[
covar(X_i, X_j) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(X_j - \bar{X})}{n-1} \quad \text{(C.6)}
\]

In which the covariance of the data is a measure of the joint variability of two variables. In our case this is a measure of the joint variability between the spectral intensity at two different wavenumbers (represented by \( X_i \) and \( X_j \)). There is large amount of data collected using Raman Spectroscopy, for the lung study in Chapter 3, we collected over 1300 wavenumbers. This means that over 1300 spectral intensities (variables) were collected with each measurement. The most convenient and straight forward way to calculate all the covariance values collected with Raman Spectroscopy is through matrix algebra. The covariance of each variable combination can be calculated and easily visualized, shown in equation C.8.
\[ C_{n \times n} = (C_{i,j}, C_{i,j} = \text{cov}(\text{Dim}_i, \text{Dim}_j)) \]  

Where \( C_{n \times n} \) is a matrix with \( n \) rows and \( n \) columns where \( \text{Dim}_i \) and \( \text{Dim}_j \) are the \( i \)th and \( j \)th dimension of the matrix respectively. In other words, a square matrix exists \((n \times n)\) where each entry in the matrix is the result of calculating the covariance between two variables. This can be visualized more easily in equation C.9.

\[
C_{n \times n} = C = \begin{bmatrix}
\text{cov}(X_1X_1) & \text{cov}(X_1X_2) & \cdots & \text{cov}(X_1X_j) \\
\text{cov}(X_2X_1) & \text{cov}(X_2X_2) & \cdots & \text{cov}(X_2X_j) \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}(X_iX_1) & \text{cov}(X_iX_2) & \cdots & \text{cov}(X_iX_j)
\end{bmatrix} \tag{C.9}
\]

Further there exists a linear combination of the variables, shown in equation C.10, in which a combination of the \( X_i \) values will predict the outcome \( Y_i \).

\[
Y_1 = e_{1,1}X_1 + e_{1,2}X_2 + \cdots + e_{1,p}X_p \\
Y_2 = e_{2,1}X_1 + e_{2,2}X_2 + \cdots + e_{2,p}X_p \\
\vdots \quad \vdots \quad \vdots \\
Y_p = e_{p,1}X_1 + e_{p,2}X_2 + \cdots + e_{p,p}X_p \tag{C.10}
\]

Where, similar to a linear regression, \( Y_i \) is being predicted through a combination of \( X_i \) values where \( e_{ij} \) values are regression coefficients. Therefore the variance of \( Y \) variables can also be calculated, as shown in equation C.11.
\[ \text{var}(Y_i) = \sum_{k=1}^{p} \sum_{l=1}^{p} e_{ik} e_{il} \text{cov}(X_k, X_l) \quad (\text{C.11}) \]

Therefore, the first principal component is the linear combination (among all combinations) of X variables that have the maximum variance. In our case, this is the weighted linear combination of spectral intensities at a specific wavenumber that accounts for the most variance among the data. In other words, the coefficients \( e_1, 1, e_1, 2, \ldots, e_1, p \) are selected so the variance (equation C.11) is maximized. This is subject to the constraint that the sum of the squared coefficients is equal to one. The second principal component is the linear combination (among all remaining combinations) of X variables that accounts for the maximum amount of the remaining variation. In other words, the coefficients \( e_2, 1, e_2, 2, \ldots, e_2, p \) are selected to maximize the remaining variance. This is subject to the constraint that the covariance between the first and second component is zero. Further principal components are calculated in the same manner. These ‘e’ coefficients are calculated by determining the eigenvector and eigenvalues of the covariance matrix C (equation C.9), and most statistical software packages are equipped with this feature. These principal component coefficients are also called loadings in the MATLAB Software (version 2013b, Math-Works). The PC factors, or scores, are the representation of the variables X in the new principal component space. In other words, we can project X onto the principal components to transform the data.

**Dimension Reduction via PLS**

For PLS analysis, we followed the methodology in (99), which is similar to the PCA methodology. Again as discussed in Chapter 3, the training process consisted of the following procedures: (1) the mean and standard deviation of the training spectra data set were calculated
as described by equations C.4 and C.5 above, (2) each spectrum in the training data set was standardized by subtracting the mean and then dividing by the standard deviation, (3) the standardized training spectra were analyzed with NIPALS algorithm (nonlinear iterative partial least squares) with the classification of the training spectra set to the known value. The weight factors, the loadings, the regression coefficient and the factor scores of the training spectra were obtained, (4) a general discrimination model was developed from the training spectra which could be used directly to predict new cases – which is discussed below.

PLS models are based on principal components. Similar to the PCA discussed above, the loadings and factors are also calculated for the independent data set (in our case the spectral intensities at each wavenumber). Yet, PLS models will also calculate the principal components for the dependent data set (in our case the histopathological grouping of each spectra) and determine a regression equation between the two sets of principal components (not the original data) that maximizes the covariance between groupings (histopathological groups). In other words, the goal of PLS is to decompose both the input and response matrices such that the covariance between groups is maximized (equation C. 12) and variance within a group is minimized.

$$\max (cov (X_w, Y)) = \max \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$  (C.12)

Where $X_w$ is each value of the independent data set (spectral intensities) and $Y$ is each value of the independent data set (histopathological group). The PLS factors and loadings can also be calculated from the Singular Value Decomposition (SVD) of the covariance matrix, but
this method cannot account for missing data (140). We therefore calculated PLS factors through the more robust NIPALS method which is an iterative process and can be seen below (for simplicity matrix notation is used) (137, 141, 142).

1) \( W = X'Y \)  
   W is a unit vector that maximizes \( \text{Cov} (X_w, Y) \)

2) \( W = W/ \|W\| \)  
   Normalize this unit vector to have length of 1. W is the weight factor.

3) \( T = XW \)  
   T is the new latent feature of X. Therefore, these are the factor scores of X.

4) \( C = Y'T/T'T \)  
   This is the linear regression coefficient of Y as a function of T

5) \( C = C/ \|C\| \)  
   Normalize this unit vector to have length 1.

6) \( U = YC \)  
   U is the new latent feature of Y. Therefore, these are the factor scores of Y.

7) Repeat until there is convergence

8) \( P = X'T/T'T \)  
   P is the projected X onto the new latent feature. These are the factor loadings of X.

9) \( Q = Y'U/U'U \)  
   Q is the projected Y onto the new latent feature. These are the factor loadings of Y.

10) \( E = X - TP' \)  
    Which is the residuals of X that are unaccounted for by TP’

11) \( F = Y-TC \)  
    These are the residuals of Y that are unaccounted for by TC. This is the intercept.

The weight factors (W), the loadings (P and Q), the regression coefficient (C) and the factor scores (T and U) of the training spectra were then used to develop a regression model which could be used to predict new cases using a classification model.
Classification via GDA

1) PCA as a Dimension Reduction:

Once the PC factors of the training cases and PC loadings were obtained, a generalized (linear) discrimination (GDA) model was developed which could be used directly to predict the new cases, shown in equation C.13.

\[
Y = b_0 + b_1 PC_1 + b_2 PC_2 + \cdots + b_k PC_k
\]  

(C.13)

Where \( b_i \) are coefficients in the linear combination, the PC values are the PC factors (scores) determined through the above PCA and \( Y \) is the histopathological group value. The training data allows for the known value of all variables except for the \( b_i \) coefficients. Once the \( b_i \) coefficients are determined (143-145) these can be used to calculate the predicted \( Y \) value (group) of an unknown spectrum.

The testing process consisted of the following procedures: (1) the test spectrum was standardized by removing the mean and dividing by the standard deviation obtained from the training spectra, (2) the PC factors of the test spectrum was calculated based on the PC loadings from the training spectra, and (3) a posterior probability of the testing spectrum was obtained from the discrimination model developed in the training procedure. The above procedures were repeated until all the spectra were left out once (and only once) for testing.
2) PLS as a Dimension Reduction

Once the PLS factors of the training cases and PLS loadings were obtained, a discrimination (DA) model was developed which could be used directly to predict the new cases, shown in equation C.14 (continued with matrix notation).

\[ Y = XB + F \]  \hspace{1cm} (C.14)

Where \( Y \) is the histopathological group value, \( X \) is the spectral intensities, \( B \) is the regression coefficient and \( F \) is the intercept (step 11 NIPALS method). \( B \) can be calculated from equation C.15.

\[ B = WC' \]  \hspace{1cm} (C.15)

Where \( W \) is the weight factor found through NIPALS (step 2) and \( C \) is the regression coefficient found through NIPALS (step 4).

The testing process consisted of the following procedures: (1) the test spectrum was standardized by removing the mean and dividing by the standard deviation obtained from the training spectra, (2) the factor scores of the test spectrum were calculated based on the weight factors from the training spectra, and (3) A posterior probability of the testing spectrum was obtained from the discrimination model developed in the training procedure. The above procedures were repeated until all the spectra were left out once (and only once) for testing.

Posterior Probability and ROC

The posterior probability of the test spectrum was calculated through Bayes theorem and is shown in equation C.16.
\[ PP = \frac{g_1(y)}{g_1(y) + g_2(y)} \]  \hspace{1cm} (C.16)

Where \( g_1(y) \) is the predicted histological value of a test spectrum and \( g_2(y) \) are the probability densities of each histopathological grouping. These are Gaussian and given by equation C.17.

\[ g_i(y) = \frac{1}{\bar{z}_i \sqrt{2\pi}} e^{-\frac{(y-\bar{\mu}_i)^2}{2\bar{z}_i^2}} \]  \hspace{1cm} (C.17)

Where \( \bar{z}_i \) is the standard deviation of the histopathological group and \( \bar{\mu}_i \) is the mean of the group. A LOO-CV was performed to treat each spectrum as a test spectrum. Once this was complete, each spectrum had an associated PP value of which an ROC can be generated from. This is generated by creating a grouping or class cut line through the PP values and counting which ones are on the correct side of the line. In other words, determining the sensitivity and specificity of the PP values based on that cut line. The line can be moved sequentially through the data in order to find an optimal location that gives the best sensitivity and specificity (classification) of the data.