Using Raman Spectroscopy for Intraoperative Margin Analysis in Breast Conserving Surgery

Submitted by **Thomas James Eliot Hubbard** to the

University of Exeter

as a

thesis for the degree of

Doctor of Medicine

in September 2020

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signature: ……  ……..

Abstract

Breast Conserving Surgery (BCS) in the treatment of breast cancer aims to provide optimal oncological results, with minimal tissue excision to optimise cosmetic outcome. Positive margins due to an inadequate resection occurs in 17% of UK patients undergoing BCS and prompts recommendation for further tissue re-excision to reduce recurrence risk. A second operation causes patient anxiety and significant healthcare costs. This issue could be resolved with accurate intra-operative margin analysis (IMA) to enable excision of all cancerous tissue at the index procedure. High wavenumber Raman Spectroscopy (HWN RS) is a vibrational spectroscopy highly sensitive to changes in protein/lipid environment and water content –biochemical differences found between tumour and normal breast tissue. We proposed that HWN RS could be used to differentiate between tumour and non-tumour breast tissue with a view to future IMA.

This thesis presents the development of a Raman system to measure the HWN region capable of accurately detecting changes in protein, lipid and water content, in the presence of highly fluorescent surgical pigments such as blue dye that are present in surgically excised specimens. We investigate the relationship between changes in the HWN spectra with changes in water content in constructed breast phantoms to mimic protein and lipid rich environments and biological tissue. Human breast tissue of paired tumour and non-tumour samples were then measured and analysed. We found that breast tumour tissue is a protein rich, high water, low fat environment and that non-tumour is a low protein, fat rich environment with a low water content, and this can be used to identify breast cancer using HWN RS with excellent accuracy of over 90%.

This thesis demonstrates a HWN RS Raman system capable of differentiating between tumour and non-tumour tissue in human breast tissue, and this has the potential to provide IMA in BCS.

Acknowledgements

I would like to thank my three supervisors – Prof Nick Stone, Prof Angela Shore and Mr Douglas Ferguson who have helped me so much over the last few years, from conception and project development to completion. Angela has provided close, detailed supervision of the project and I appreciate the eye for specifics to keep the project in check. I also appreciate the huge amount of moral support as well as the time and effort she has given me. Mr Ferguson has been an excellent clinical supervisor, facilitating the clinical aspect of the project, helping to work around the multiple logistical problems with organising a clinical research project, and providing support and career guidance. Particular thanks goes to Nick, who's vision and experience I have drawn on throughout. I feel he is the ideal supervisor, guiding where necessary, encouraging me to problem solve and better myself, but also being hands on with solving issues when necessary and giving superlative detailed supervision. I appreciate the significant amount of time and commitment he has given to the project. The combination of supervisors has led to an incredibly interesting few years of scientific inquiry, and intellectual stimulation for me personally, and an excellent research project with corresponding academic outcomes.

Although there were difficulties with initiating the hospital project with logistics, and then the Covid pandemic, a number of people helped me in the planning of the clinical phase. Claire Murray (pathologist) was fantastic in helping to develop the handling of pathological specimens, and her time spent on this, and on the second reading of H+E slides is appreciated. Natalie (breast secretary team leader) and Pauline (research nurse) also helped immensely with planning the project, answering queries, and identifying participants suitable for study inclusion.

I would like to thank the biophysics group for accommodating a clinician into their fold. I have found the discussions relating to a wide range of matters, and their insight into academia very useful. Particular thanks goes to Adrian for involving me to collaborate with him on an interesting paper. Alex Dudgeon has been a great help in the last year of the project, and I appreciate the work he did in obtaining the frozen breast samples that we analysed together, and his subsequent Matlab tutelage. A huge thank you to Ben Gardner for all of his help,

from helping me clean up broken slides (!), to navigating hospital SOPs, to Matlab queries, your time and considerable knowledge has been greatly appreciated.
Finally I would like to thank my family for their help and support during these two years of immense hard work, which have enabled me to complete this thesis. My parents continue to be emotionally and practicably supportive of all my endeavours, which is a privilege. To Jo and Eve, an extraordinary couple of years, thank you for everything.

Table of Contents

TABLE OF FIGURES **11**

LIST OF TABLES **17**

ABBREVIATIONS **18**

CHAPTER 1 : **BACKGROUND** **21**

1.1 BREAST CANCER 21

1.1.1 Incidence *21*

1.1.2 Anatomy of the breast *24*

1.1.3 Pathology of breast cancer *27*

1.1.4 Diagnosis *35*

1.1.5 Management of early breast cancer *39*

1.1.6 Summary *44*

1.2 BACKGROUND OF RAMAN SPECTROSCOPY 45

1.2.1 The theory of Raman scattering *45*

1.2.2 Experimental measurement of Raman scattering *50*

1.2.3 Summary *53*

1.3 RAMAN SPECTROSCOPY FOR RAPID INTRA-OPERATIVE MARGIN ANALYSIS OF SURGICALLY EXCISED TUMOUR SPECIMENS 54

1.3.1 Introduction *54*

1.3.2 Current methods of margin analysis *55*

1.3.3 Methods of Intra-operative Margin Analysis (IMA) *56*

1.3.4 Raman spectroscopy *59*

1.3.5 The use of RS to determine the surgical margin *64*

1.3.6 Translating Raman to the clinical environment for IMA *80*

1.3.7 Summary *84*

1.4 THE ROLE OF WATER CONTENT IN BREAST CANCER DIAGNOSIS 85

1.4.1 Introduction *85*

1.4.2 Theories to explain differences in water content in cancerous tissue *85*

1.4.3 Water content in normal breast tissue *86*

1.4.4 Differentiating normal from cancerous tissue using water in breast cancer *88*

1.4.5 Differentiating normal from cancerous tissue using water with HWN RS in other cancers *89*

1.4.6 Differentiating normal from cancerous tissue using water with HWN RS in breast cancer *91*

1.4.7 Summary *93*

CHAPTER 2 : **DEVELOPING A RAMAN SYSTEM FOR INTRAOPERATIVE MARGIN ANALYSIS** **95**

2.1 INTRODUCTION 95

2.2 MATERIALS AND METHODS 96
2.2.1 Measured materials 96
2.2.2 Raman Spectroscopy equipment configuration and spectral acquisition 104
2.2.3 Calibration 107
2.2.4 Data recording and Analysis 111
2.3 EVALUATING BREAST PHANTOMS AND DATA ANALYSIS PROCEDURES 116
2.4 EVALUATION OF RAMAN SYSTEMS 121
2.4.1 NP1- 680 nm laser excitation 121
2.4.2 NP-2 785 nm set up 125
2.4.3 NP-3 785 nm InGaAs set up 133
2.5 DISCUSSION 138
2.6 CONCLUSION 141

CHAPTER 3 : **INVESTIGATING THE ABILITY OF NP3 RAMAN SYSTEM TO QUANTIFY CHANGES IN WATER CONTENT USING HIGH WAVENUMBER RAMAN SPECTROSCOPY 143**

3.1 INTRODUCTION 143
3.2 METHODS 145
3.2.1 Establishing the optimal signal acquisition time 145
3.2.2 Phantom Production and Measurement 150
3.2.3 UV Vis experiments 153
3.3 RESULTS 153
3.3.1 Gelatine phantoms 153
3.3.2 Soya bean oil phantoms 155
3.3.3 Pork phantoms 159
3.3.4 Mixed protein/lipid phantoms 163
3.4 DISCUSSION 166
3.4.1 NP3 can measure changes in water concentration at physiological concentrations 166
3.4.2 The effect of changes in water concentration on the HWN spectrum in three different environments 168
3.4.3 The water/total area ratio relationship with water concentration is dependent on protein/lipid environment 177
3.5 CONCLUSION 180

CHAPTER 4 : **HIGH WAVENUMBER RAMAN SPECTROSCOPY FOR THE IDENTIFICATION OF BREAST CANCER IN FRESH FROZEN SAMPLES 183**

4.1 INTRODUCTION 183
4.2 METHODS 185
4.2.1 Human Tissue samples 185
4.2.2 Needle Probe Raman Spectroscopy configuration 186
4.2.3 Fresh Frozen sample Breast Tissue measurements 186

4.2.4 Estimation of water content by dehydration 186
4.2.5 Data Processing 187
4.2.6 Data analysis 195
4.2.7 Raman micro spectrometry experiments 198
4.3 RESULTS 200
4.3.1 Patient/ Tumour demographics 200
4.3.2 Fresh frozen specimen measurements 201
4.3.3 Estimation of water content by dehydration 252
4.3.4 Raman micro spectrometry experiments 255
4.4 DISCUSSION 261
4.4.1 Interpretation of results 261
4.4.2 Limitations of methods 275
4.4.3 Tissue types presenting a diagnostic challenge 279
4.5 CONCLUSION 281

CHAPTER 5 : **AN OPTIMISED RAMAN SYSTEM FOR INTRAOPERATIVE MARGIN ANALYSIS 283**

5.1 INTRODUCTION 283
5.2 METHODS 284
5.2.1 HH Raman system configuration 284
5.2.2 Breast Phantom construction 285
5.2.3 Measurement protocol 286
5.2.4 Data processing and analysis 286
5.3 EXPERIMENTAL PROTOCOL AND RESULTS 287
5.3.1 Assessment of probe detecting changes in water content 287
5.3.2 The effect of changing measurement parameters on signal acquisition 289
5.3.3 Assessment of probe in 'tumour' detection 301
5.4 DISCUSSION 308
5.4.1 Clinical context of findings 308
5.4.2 Limitations 311
5.4.3 Further work 312
5.5 CONCLUSION 313

CHAPTER 6 : **DISCUSSION 314**

6.1 SUMMARY OF THESIS FINDINGS 314
6.2 CLINICAL APPLICATIONS OF THESIS FINDINGS 316
6.3 COMPARING FINDINGS TO OTHER METHODS OF IMA 318
6.4 LIMITATIONS 323

CHAPTER 7 : **FURTHER WORK AND CONCLUSION 327**

7.1 FURTHER WORK 327
7.1.1 Validation of diagnostic findings in fresh breast tissue 327

7.1.2 Probe optimisation 328
7.1.3 Data analysis optimisation 329
7.1.4 Clinical trial 329
7.2 CONCLUSION 330
APPENDICES 331
REFERENCES 335

Table of Figures

Figure 1-1 Summary of NICE guidelines for those at greater risk of breast cancer based on Family History. 22

Figure 1-2. Graph showing Age-Standardised Ten-Year Net Survival, Women (Aged 15-99), England and Wales, 1971-2011 23

Figure 1-3 Diagram showing the underlying structures of the breast, its anatomical relationships, the arterial supply and venous and lymphatic drainage. 25

Figure 1-4 Diagram showing Cross – section of the female breast demonstrating the relationship of fat, lobules and ducts. 27

Figure 1-5 Diagram showing the difference between normal breast duct, DCIS and invasive breast cancer. 29

Figure 1-6. Kaplan Meier survival graph of overall survival for 72 patients with locally advanced breast disease in a Norway cohort according to molecular subtype. 32

Figure 1-7 TNM staging classification and stage grouping. 34

Figure 1-8 Summary box outlining the 'fast track' referral criteria from primary care to secondary care for urgent evaluation by a Breast specialist[41].....36

Figure 1-9 Diagram demonstrating options for incisions for Breast conserving surgery. 42

Figure 1-10 The Electromagnetic spectrum on the wavelength scale in metres with associated Energy per photon in electron volts (eV) 45

Figure 1-11. A Jablonski energy diagram of quantum energy transitions for Rayleigh and Raman scattering. 47

Figure 1-12 Schematic diagram demonstrating the different modes of bond vibrations that may occur within a molecule.....48

Figure 1-13 Representative Raman spectra demonstrating the cellular molecular components that can be measured using high wavenumber Raman Spectroscopy. 50

Figure 1-14 Diagram demonstrating a typical set up for a system to perform Raman spectroscopy. 51

Figure 1-15 A graphic to illustrate the concept of tumour margins. 55

Figure 1-16 A graphical representation of the ideal paradigm for the surgical workflow of intraoperative margin analysis (IMA) by Raman Spectroscopy. 59

Figure 1-17 Examples of fibre optic probes capable of Raman Spectroscopy measurement in a clinical setting for cancer diagnosis, or application to in vivo surgical guidance to provide IMA, images Authors own.....62

Figure 1-18 Example of assigning Raman spectra to tissue structures and ductal carcinoma of breast tissue to inform the diagnostic algorithm. 67

Figure 1-19 Examples of Raman probe systems that could be used for IMA using Raman spectroscopy. 83

Figure 1-20 Figure of Raman spectra showing the mean spectra of normal and cancer tissue in a selected HWN region for Ex vivo biopsy for normal (EVNT) and cancerous (EVCT) breast tissue from rats. .92

Figure 1-21. HWN Raman spectra of normal (non-cancerous), and cancerous human breast tissue (infiltrating ductral carcinoma) compared to the bulk neat water in bulk tissue. 93
Figure 2-1 HWN Raman spectra from a set of gelatine phantoms of different water concentrations. 98
Figure 2-2 Photograph of the pure gelatine phantoms at decreasing water concentration 99
Figure 2-3 Photograph of complex gelatine phantoms with increasing concentrations of intralipid 100
Figure 2-4 Photograph of pork fat (left) and meat (right) that have been stained with 1% patent blue dye. 102
Figure 2-5 Pork meat that has been burnt with some carbonisation evident.....103
Figure 2-6 Pork meat stained with Haemoglobin solution 103
Figure 2-7 A picture of the needle component of the Raman needle probe which delivers and collects light to the specimen for Raman analysis. 105
Figure 2-8 Schematic diagram showing the Raman needle probe set up, modified from ref [164]. 106
Figure 2-9 Bar chart of maximum Raman spectral intensity at 3390 cm^{-1} at different water concentrations calculated by BSA/water mixtures. 113
Figure 2-10 Graph showing Raman spectra in the HWN region of different concentrations of BSA/water solutions.....117
Figure 2-11 Graphs showing calculation of the water/total area ratio in BSA/water solutions with the InVia spectrometer. 117
Figure 2-12 Graph showing gelatine phantoms of the same water concentrations in 3 different experiments and the water /total ratio. 120
Figure 2-13 Graph showing the same gelatine phantom measured at time zero and 4 hours later. 120
Figure 2-14 Raman spectra of a range of varying water concentrations of gelatine phantoms measured using NP1 Raman system. 122
Figure 2-15 Using NP1 to measuring water/total area ratio in gelatine phantoms 122
Figure 2-16 Demonstrating the ability to measure changes in water concentration in complex intralipid phantoms measured with NP1. 123
Figure 2-17 Results of measurements of pork phantoms with NP1. 123
Figure 2-18 Raman spectra of a range of gelatine phantoms of varying water concentrations measured using NP2 Raman system. 126
Figure 2-19 Using NP2 to measure water/total area ratio in gelatine phantoms 126
Figure 2-20 Demonstrating the ability to measure changes in water concentration in complex intralipid phantoms measured with NP2. 127
Figure 2-21 Results of measurements of pork phantoms with NP2 system. 128
Figure 2-22 Results of measurement of blue dye (BD) gelatine phantoms with NP2. 129
Figure 2-23 Quantum efficiency of the Princeton Instruments camera range versus wavelength. 131
Figure 2-24 Quantam efficiency at 20°C of the Andor InGaAs camera range versus wavelength which is shown by the continuous red line. Modified from reference [339] 132
Figure 2-25 Raman spectra of a range of varying water concentrations of gelatine phantoms measured using NP1 Raman system. 133
Figure 2-26 Using NP3 to measure water/total area ratio in gelatine phantoms 134

Figure 2-27 Demonstrating the ability to measure changes in water concentration in complex intralipid phantoms measured with NP3. 134
Figure 2-28 Results of measurements of pork phantoms with NP3 system. 135
Figure 2-29 Results of measurement of blue dye gelatine phantoms with NP3. 136
Figure 3-1 – Spectrum showing the Dark background and Smoothed Dark background at 1 second acquisition time with 5 accumulations. 146
Figure 3-2 Figure showing the effect of Dark noise baseline subtraction on a HWN spectrum. 146
Figure 3-3 – Figure showing the difference in signal to noise in Raman spectra obtained with two different measurement protocols, both with an overall measurement time of 5 seconds. 148
Figure 3-4- Figure showing the difference in signal to noise in Raman spectra obtained with three different measurement protocols, all with an overall measurement time of 25 seconds. 149
Figure 3-5 Figure showing Raman spectra in the HWN region of different water concentrations of gelatine phantoms and the relationship with water/total area ratio 154
Figure 3-6 Graphs demonstrating the changes in the HWN region with changes in water concentration in a protein only environment. 155
Figure 3-7. Figure showing Raman spectra in the HWN region of different water concentrations of soya bean oil phantoms and the relationship with water/total area ratio 156
Figure 3-8 – Graphs demonstrating changes in the HWN region with changes in water concentration in a lipid only environment. 157
Figure 3-9- Graph demonstrating the UV Vis absorption spectra of water, Triton X100 and soya bean oil 158
Figure 3-10 Figure showing Raman spectra in the HWN region of pork meat at different stages of dehydration and the relationship with water/total area ratio 159
Figure 3-11 – Graphs demonstrating changes in the HWN region with changes in water concentration in a biological system. 160
Figure 3-12. Graphs showing the change in Unbound versus Bound water in serially dehydrated pork. 162
Figure 3-13 Figure showing Raman spectra in the HWN region of pork fat at different stages of dehydration and the relationship with water/total area ratio 163
Figure 3-14 - Figures demonstrating the effect of a change in lipid/protein environment on the HWN spectrum in phantoms with a water concentration of 70% 164
Figure 3-15 - Figures demonstrating the effect of a change in lipid/protein environment on the HWN spectrum in phantoms with a water concentration of 80% 165
Figure 3-16 Graph showing linear relationship between known water fraction and water/total area ratio 178
Figure 4-1 High wavenumber spectra of human breast tissue using different pre-processing methods 189
Figure 4-2 Demonstrating the complex fluorescent background baseline of breast tissue specimens ...190
Figure 4-3 Demonstrating different methods of pre-processing a representative tumour spectrum 191
Figure 4-4 Demonstrating pre-processing tumour spectrum with a 6^{th} order polynomial flourescence baseline subtraction prior to min/max normalisation. 193

Figure 4-5 Representative baselined and normalised Raman spectrum obtained from a single non-tumour specimen, with demonstration of the different spectral regions measured 194
Figure 4-6 Graphs showing spectral data obtained from all measurements (n=1335) after pre-processing 201
Figure 4-7 Mean HWN spectra of all specimens (n=1335) according to diagnosis demonstrating changes in water content between tumour and non-tumour specimen spectra 202
Figure 4-8 Bar charts illustrating the difference in water/total area ratio between tumour (n=96) and non-tumour specimens (n=96) 203
Figure 4-9 ROC curve for classification by binomial logistic regression of water/total area ratio prediction of tumour versus non-tumour of all spectra (n=1335) 204
Figure 4-10 ROC curve for classification by binomial logistic regression of water/total area ratio prediction of tumour versus non-tumour of mean spectra (n=192) 205
Figure 4-11 Confusion matrices of specimen diagnosis after voting threshold analysis of water/total area ratio using the optimal threshold of water/total area ratio of 0.75 207
Figure 4-12 Confusion matrix showing the diagnostic results of the W/TAR difference cut off= 0.1 for classifying tumour versus non-tumour spectra 210
Figure 4-13 Results of binomial logistic regression model with k folds 5 cross validation for predicting tumour versus non-tumour specimen using the W/TAR difference 211
Figure 4-14 Bar chart of sample number (n=96) and the mean W/TAR Difference for the tumour specimen for that sample 212
Figure 4-15 Principal Component Analysis of all HWN spectral data (n=1335) 214
Figure 4-16 Principal components accounting for the majority of spectral variance between tumour versus non-tumour tissue, and spectral assignation 215
Figure 4-17 Results of all HWN spectra (n=1335) to compare spectra from tumour versus non- tumour specimens with PCA fed LDA analysis 216
Figure 4-18 Principal Component Analysis of mean HWN spectral data (n=192) 217
Figure 4-19 Results from spectral analysis of the mean HWN spectra for each specimen (n=192) 218
Figure 4-20 Mean Fingerprint spectra of all specimens (n=1335) according to diagnosis demonstrating differences between tumour and non-tumour specimen spectra 219
Figure 4-21 Spectral assignation of Fingerprint spectra in breast tissue 220
Figure 4-22 Principal Component Analysis of Fingerprint region spectra of all spectra (n=1335) 222
Figure 4-23 Spectral assignation in the two most significant principal components in the Fingerprint region in differentiating tumour from non-tumour tissue 223
Figure 4-24 PCA fed LDA analysis of all Fingerprint region spectral data (n=1335) to differentiate between tumour and non-tumour breast spectra 225
Figure 4-25 Principal Component Analysis of Fingerprint region spectra of mean spectra (n=192) 226
Figure 4-26 PCA fed LDA analysis of mean Fingerprint spectra (n=192) to differentiate tumour and non-tumour specimens 228
Figure 4-27 Principal component analysis of all concatenated spectral data (n=1335) 230
Figure 4-28 PCA fed LDA analysis of all concatenated spectral data (n=1335) 231

Figure 4-29 Principal Component Analysis of concatenated spectra of mean spectra (n=192).....232
Figure 4-30 Analysis results of mean concatenated spectra for each specimen (n=192) 233
Figure 4-31 Results of mean tumour specimen (n=96) spectral analysis of ER status (used as a representative example).....235
Figure 4-32 Principal component analysis of mean HWN tumour spectra according to pathological subtype classification (n=96) 237
Figure 4-33 Principal component analysis of mean FP tumour spectra specimens according to pathological subtype classification (n=96) 238
Figure 4-34 Principal component analysis of mean concatenated tumour spectra according to pathological subtype classification (n=96) 240
Figure 4-35 Mean HWN spectra of all specimens (spectra n=42, patients n=3) according to diagnosis demonstrating changes in water content between DCIS only and non-tumour specimen spectra 245
Figure 4-36 Principal Component Analysis of concatenated spectral data for patients with DCIS only (n=42) 246
Figure 4-37 Results of mean spectral analysis of non-tumour breast specimens classified according to menopausal status (n=89) 248
Figure 4-38 Results of spectral analysis of non-tumour breast specimens classified according to Age (cut off 50) (n=95) 250
Figure 4-39 Results of serial dehydration of tumour specimens (n=3) and corresponding water/total area ratio.....253
Figure 4-40 Results of serial dehydration of a non-tumour specimen with corresponding Raman spectra (n=1).....254
Figure 4-41 H+E slides of non-tumour specimens.....256
Figure 4-42 HWN Micro-Raman analysis of non-tumour specimen from patient 105.....257
Figure 4-43 H+E slides of tumour specimens 258
Figure 4-44 HWN Micro-Raman analysis of tumour specimen from patient 105 259
Figure 4-45 HWN Micro-Raman analysis of tumour specimen from patient 14 260
Figure 4-46 A schematic representation of how the suggested data processing algorithm could be used on a specimen resected to give a clinical outcome 274
Figure 5-1 HH Raman system configuration.....285
Figure 5-2 Showing Raman spectrum of pork meat and pork fat obtained with HH Raman system. 287
Figure 5-3 showing Raman spectra in the HWN region of different water concentrations of gelatine phantoms and the relationship with the water/total area ratio 288
Figure 5-4 Results of measurement of pork meat with HH Raman system with different laser excitation powers 290
Figure 5-5 Raman spectra of pork meat obtained with HH Raman system with a range of signal acquisition times (5 second – 0.1 second) 293
Figure 5-6 Raman spectra of pork meat obtained with HH Raman system with a range of probe to specimen distances (0 mm – 10 mm) 296

Figure 5-7 Raman spectra of pork meat obtained with HH Raman system with a range of depths of 90% water gelatine phantom overlying the fat (0 mm – 7 mm) 299
Figure 5-8 Raman spectra of pork meat and lard covering pork meat obtained with HH Raman system 300
Figure 5-9 Figure showing the lumpectomy phantoms.....302
Figure 5-10 Raman spectra of lumpectomy phantoms obtained with HH Raman system with a range of pork meat (to mimic 'tumour') sizes (0 mm^2 – 5 mm^2).....306
Figure 5-11 demonstrating the concept of a rapid line scan of a lumpectomy phantom 307

List of Tables

Table 2-1 Table showing the constituent parts of the complex intralipid/gelatine phantoms 100
Table 2-2 Table demonstrating the Raman shift references used for calibration of the Raman shift axis, and the laser wavelength for which they were utilised. 110
Table 2-3 Table demonstrating peak assignation for the HWN Raman spectroscopy CH stretch region in a variety of human tissues 112
Table 2-4 Peak assignation table for the HWN Raman spectroscopy OH stretch region 112
Table 2-5 Table showing the validation of using the water/total area ratio to calculate water concentration. 118
Table 2-6 Table showing the actual water concentration of intralipid phantoms, the calculated water/total area ratio, and the predicted water concentration based on the BSA/water calibration. 119
Table 7 - Table to show the composition of mixed phantoms composed of protein, lipid and water 152
Table 4-1 Table of patient and tumour characteristics for the samples that were measured in this study. 200
Table 4-2 Summary table of results following Model predictions of categorising spectra as tumour versus non-tumour based on water/total area ratio in all spectra (n=1335) 204
Table 4-3 Summary table of results following Model predictions of categorising spectra as tumour versus non-tumour based on water/total area ratio in mean spectra (n=192) 206
Table 4-4 Table demonstrating the change in diagnostic accuracy with a change in water/total area ratio threshold (Ratio Threshold) to classify the majority of a specimen using voting threshold 1. 208
Table 4-5 Peak assignment of significant Principal component scores in fingerprint spectral analysis...224
Table 4-6 Peak assignment of significant Principal component scores in fingerprint mean spectral analysis Biochemical peak assignation is according to the literature [103] [356] 227
Table 4-7 Summary statistics table of best performing spectral analysis and water/total area ratio analysis 234
Table 4-8 Results of ability of the three different spectral regions to predict pathological subtype by PCA fed LDA analysis performed on mean spectral data (n=96) 241
Table 4-9 Table demonstrating analysis of mean spectra of tumour specimens (n=96) of pathology subtypes 243
Table 4-10 Summary results table of analysis of non-tumour specimens mean spectra classified by physiological characteristics 251
Table 6-1 Table demonstrating the sensitivity (Sens.) and specificity (spec.) of commonly used Intraoperative Margin Analysis (IMA) techniques 319
Table 6-2 Summary table of diagnostic accuracy of different techniques based on Raman spectroscopy for differentiating normal breast tissue from breast tumour. 322

Abbreviations

ANOVA	Analysis of Variance
A.U.	Arbitrary Units
AUC	Area under the Curve
BCC	Basal Cell Carcinoma
BCS	Breast Conserving Surgery
BMI	Body Mass Index
BRCA	BReast CAncer gene
BSA	Bovine Serum Albumin
CARS	Coherent anti-Stokes Raman Spectroscopy
CCD	Charge – Coupled Device
DCIS	Ductal Carcinoma *in Situ*
DNA	Deoxyribose nucleic acid
DOS	Diffuse Optical Spectroscopy
DRS	Diffuse Reflectance Spectroscopy
ECOC	Error – correcting Output Codes
ER	(o)Estrogen Receptor
FP	Fingerprint
Hb	Haemoglobin
H + E	Haemolysin and Eosin
HER2	Human Epidermal Receptor 2
HH Raman system	Handheld Raman system
HWN	High Wavenumber
HWN RS	High Wavenumber Raman Spectroscopy
IDC	Invasive Ductal Carcinoma
IMA	Intraoperative Margin Analysis
InGaAs	Indium Gallium Arsenide
LCIS	Lobular Carcinoma *in situ*
LDA	Linear Discriminant Analysis
LOOCV	Leave-one-out-cross-validation
MDT	Multi- Disciplinary Team

MRI	Magnetic Resonance Imaging
NHS	National Health Service
NICE	National Institute of Clinical Excellence
NIR	Near-Infra Red
NP3	Needle Probe 3 (Raman system)
OCT	Optical Coherence Tomography
PCA	Principal Component Analysis
PLS	Partial Least Squares
PPV	Positive Predictive Value
PR	Progesterone receptor
PTFE	Polytetraflouroethylene
QE	Quantum efficiency
ROC	Receiver Operator Characteristic
RS	Raman Spectroscopy
SCC	Squamous Cell Carcinoma
SERS	Surface -enhanced Raman Spectroscopy
US	Ultrasound
USA	United States of America
UV	Ultraviolet (light)

CHAPTER 1: Background

1.1 Breast cancer

1.1.1 Incidence

Breast cancer is the most common cancer to affect women in the world and the second most common cancer overall, accounting for 6.6% of all cancer deaths [1]. There is an incidence of 55,000 cases per year in the UK with a 1 in 8 lifetime risk for women, accounting for 15% of all cancer diagnoses in the UK and over 11,000 deaths per year [2]. This represents a significant burden to the NHS and a major public health issue for the UK. Cancer survival in the UK lags behind many other European countries, and an emphasis on earlier diagnosis and 5 other priority areas was highlighted in the Cancer Strategy Implementation Plan 2015-2020 to improve cancer survival, and demonstrates that cancer care remains a priority for NHS England [3].

The incidence rate of breast cancer is rising and there is a much higher incidence in high-income countries such as western Europe and the USA [4]. This is postulated to be due to the primary risk factors being difficult to modify, and lifestyle risk factors becoming increasingly prevalent in Western lifestyle, such as obesity, late age of first child and reduced fertility rate. The known protective factors of breast cancer are breastfeeding – the relative risk decreases by 4.3% for every 12 months of breastfeeding - and a moderate degree of physical activity [5].

The risk factors associated with breast cancer are [1]:

Age – increasing age, however the rate doubles every 10 years until menopause then the rate slows [6]. Therefore, the largest proportion of breast cancer is in women aged 45-69.

Geographical location – Higher incidence in western Europe and USA [4]. Migrants from a low risk area to a high risk area show the same incidence within one or two generations demonstrating the significance of environmental factors [6], with a corresponding change in breast cancer mortality [7].

Timing of onset/cessation of menses – early age at menarche and later age at menopause is associated with an increased risk of breast cancer

Reproduction – nulliparity, late age at first birth and having fewer children increases the risk of breast cancer

Exogenous hormone intake – oral contraception use (particularly the combined pill containing oestrogens) [8] and hormone replacement therapy

Nutrition – increased alcohol intake increases the risk of breast cancer

Cigarette Smoking – possible direct link to increased incidence, though it is not a significant risk factor [9]

Anthropometry – increased weight, weight gain during adulthood and body fat distribution

Breast density- dense breasts increase the risk of breast cancer. Although breast density is affected by age and certain hormonal agents (such as tamoxifen and HRT), the relative risk of breast cancer is increased four to fivefold in women with high density breasts compared to those with low density breasts [5].

Genetics –Being a carrier of the identified gene mutation BRCA 1 or 2 increases risk of breast cancer [10], with a lifetime risk of 60-85% for BRCA1 carriers, and 50-85% for BRCA 2 carriers [6]. 4-5% of breast cancer is due to inheritance of a high-penetrance, autosomal dominant, cancer pre-disposing gene [11]. Therefore, certain family histories, even without identified gene mutations, may result in a clinically higher risk requiring increased screening (Figure 1-1).

1.3.3 People without a personal history of breast cancer who meet the following criteria should be offered referral to secondary care:

- one first-degree female relative diagnosed with breast cancer at younger than age 40 years **or**
- one first-degree male relative diagnosed with breast cancer at any age **or**
- one first-degree relative with bilateral breast cancer where the first primary was diagnosed at younger than age 50 years **or**
- two first-degree relatives, or one first-degree and one second-degree relative, diagnosed with breast cancer at any age **or**
- one first-degree or second-degree relative diagnosed with breast cancer at any age and one first-degree or second-degree relative diagnosed with ovarian cancer at any age (one of these should be a first-degree relative) **or**
- three first-degree or second-degree relatives diagnosed with breast cancer at any age. **[2004]**

Figure 1-1 Summary of NICE guidelines for those at greater risk of breast cancer based on Family History.

Patients with a family history that fits these criteria should be referred to a secondary care family history clinic for consideration of further investigation/surveillance [12]

Although incidence is increasing, mortality rates have improved substantially over the last 40 years. Survival at 1, 5 and 10 years has improved. At 10 years, survival has risen from 40% in 1971-72 to 78 % for those patients diagnosed in 2010-11 (Figure 1-2). One reason for this is improved early diagnosis and the establishment of the UK breast screening programme in 1987; even adjusting for the lead time bias introduced by early detection, screen-detected cancers had a substantial survival advantage compared to those that were symptom detected [13]. Another reason is the substantial improvements in essentially all treatments particularly radiotherapy, hormonal therapy, molecular profiling of cancers with targeted treatments, and the advent of individualised treatments based on genetic profiling. This rate of improved long term survival has not plateaued, suggesting there is still room for improvement [14].

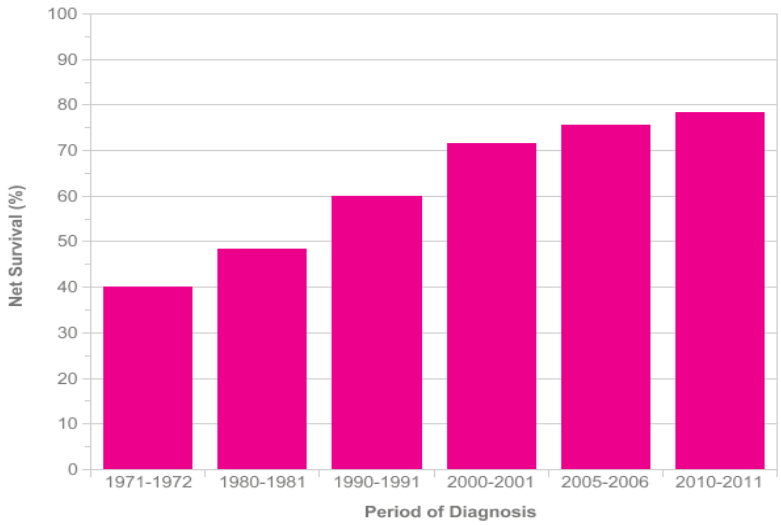


Figure 1-2. Graph showing Age-Standardised Ten-Year Net Survival, Women (Aged 15-99), England and Wales, 1971-2011 .

There is an increase in net 10 year survival in those diagnosed with Breast Cancer. Credit – Cancer Research UK [15]

The cost of treating breast cancer in the USA in 2008 is estimated to be USD $20 – 100, 000 for a lifetime per patient cost of breast cancer [16]. The cost to the whole US healthcare economy is estimated to be USD 15-20 billion for the year 2001. USD 16.5 billion is spent each year on breast cancer diagnosis and treatment, and as it occurs in a young age group, there are economic factors such as lost work productivity which increases the cost. The cost continues to rise as the

incidence increases, and treatment (particularly new, targeted drug treatment) gets more sophisticated leading to increased survivorship.

It can be seen that Breast cancer is an increasingly important public health issue, with a large number of women being diagnosed with the disease and this is a financially expensive disease for society. Improving efficiency of diagnosis and reducing treatment costs whilst continuing to see a continuation in survivorship is a priority for the NHS.

1.1.2 Anatomy of the breast

The mammary glands are modified sweat glands in the superficial fascia anterior to the pectoral muscles and the anterior thoracic wall. The breast lies on the underlying muscle separated by a layer of deep connective tissue that forms the retromammary space. An overview of breast anatomy is demonstrated in Figure 1-3.

Surface Anatomy

Although there is a degree of variation in size, the breast is positioned on the anterior thoracic wall between ribs II and VI, overlying the pectoralis major muscle. The breast extends superolaterally around the lower margin of pectoralis major and enters the axilla.

Arterial supply and venous drainage [17]

The breast gains its blood supply from the vessels of the axillary artery (superior thoracic, thoraco-acromial, lateral thoracic and subscapular arteries, medially from branches of the internal thoracic artery and perforator vessels from the thoracic wall from the second third and fourth intercostal arteries).

Venous drainage is via veins that mirror the arterial system, and drain to the axillary, internal thoracic and intercostal veins. Innervation is from the anterior and lateral cutaneous branches of the second to sixth intercostal nerves, the nipple being innervated by the fourth intercostal nerve.

Lymphatic drainage

The lymphatic system is an extensive network of channels that drain the fluid from body tissues lost from capillary beds during nutrient exchange processes

and deliver it back to the venous system [17]. The lymph system is a major route of spread of cancer from one body organ or space to another, and particularly in breast cancer, the lymph system is often the first route of invasion, therefore the anatomy of the lymph drainage is important in understanding breast cancer.

Lymphatic drainage of a breast is:

- 75% via lymphatic vessels draining laterally and superiorly into axillary nodes
- The majority of the rest drain medially to parasternal nodes associated with the internal mammary artery
- A minority drain via the lymphatics associated with the ribs [17]

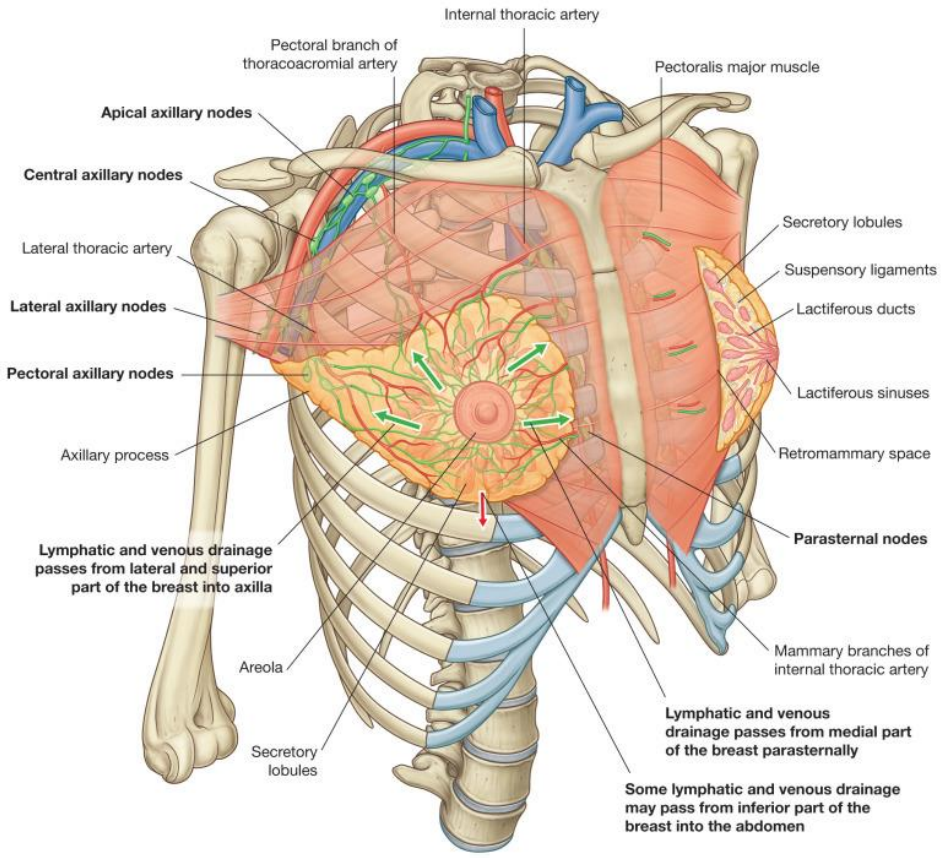


Figure 1-3 Diagram showing the underlying structures of the breast, its anatomical relationships, the arterial supply and venous and lymphatic drainage.

Lymphatic drainage is denoted by green arrows -note that Level I axillary lymph nodes are those defined below Pectoralis Minor, Level II underneath pectoralis minor, and Level III any lymph nodes above pectoralis minor. Reprinted from Gray's Anatomy for Students, 4th Ed, Drake, Vigl and Mitchell, Thorax, Page No 133., Copyright (2020), with permission from Elsevier [17]

Breast composition

The anatomy of the human breast was originally described by Cooper in 1840 based on an extensive number of dissections and the majority of the concepts remain unchanged [18]. The breast is composed of fat, and a series of ducts and associated secretory lobules that consist of clusters of alveoli containing lactocytes that during lactation undergo widespread hypertrophy and hyperplasia and synthesise breastmilk [19]. These lobules then converge to form ducts, the number of which was originally thought to be around 22, but is now thought to be a more modest 5-9 based on more recent ultrasound studies [18, 20]. These lactiferous ducts then open independently onto the nipple [17] (Figure 1-4). Adipose tissue is between lobes rather than within lobules and a connective tissue stroma surrounds the ducts and lobules, which condense in some regions to form the Astley Cooper ligaments, which support the breast. There is significant variation in the number and size of ducts and amount of glandular / fat tissue between women, but is consistent between breasts within women [20].

The development of the female breast is under hormonal control of progesterone and prolactin during pubertal mammogenesis, and continues to respond to the fluctuations of oestrogen and progesterone associated with the menstrual cycle and ageing [21]. Glandular tissue is abundant in lactation with a corresponding increase and engorgement in breast size, whereas glandular tissue reduces with age and fat increases, leading to softer and more pendulous breasts [22].

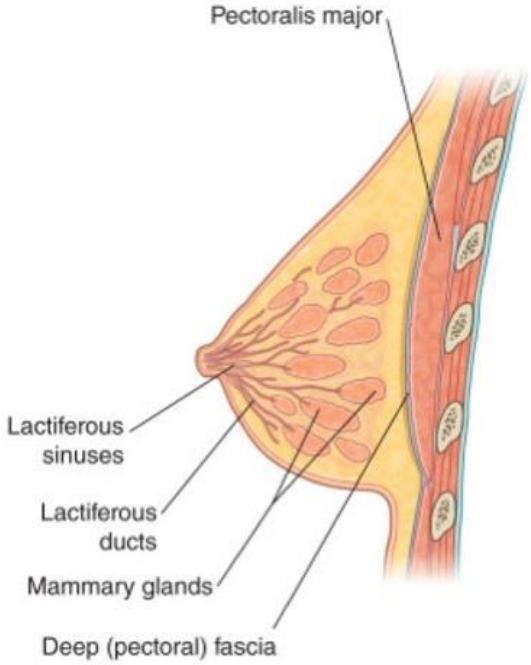


Figure 1-4 Diagram showing Cross – section of the female breast demonstrating the relationship of fat, lobules and ducts.

The mammary glands that produce milk drain to lactiferous sinuses that discharge to the nipple during lactation. Reprinted from Gray's Anatomy for Students, 4th Ed, Drake, Vigl and Mitchell, Thorax, Page No 133., Copyright (2020), with permission from Elsevier [17]

1.1.3 Pathology of breast cancer

Although there are other cancers of the breast, this thesis mainly deals with the most common types of ductal carcinoma (both *in situ* and invasive) and lobular carcinoma. The most common location of tumours within the breast are upper outer quadrant (50%), and central portion (20%) with 4% of women having bilateral primary tumours.

1.1.3.1 Carcinogenesis

The development of cancer is a result of the dysregulation of the body's normal mechanisms for controlling cellular growth. Kumar *et al.* highlight three main features that underpin carcinogenesis, and succinctly describe the fundamental shared characteristics of cancers [23]:

1. All cancers are to some degree due to DNA mutations, either induced spontaneously or through environmental insult.
- An example in breast cancer is the HER2/NEU protooncogene, that is associated with a poor prognosis
2. These genetic alterations are therefore heritable, and so family history is an important risk factor of cancer
- An example in breast cancer are the previously discussed BRCA mutations
3. Accumulation of mutations result in the hallmarks of cancer;
- Self- sufficiency in growth signals and autonomous growth which is unregulated
- lack of response to inhibitory mechanisms of cell proliferation
- evasion of cell death
- limitless replicative potential
- development of angiogenesis to sustain tumour growth
- ability to invade local tissues and systemic spread (distant metastases)
- reprogramming of metabolic pathways
- ability to evade the immune system

In breast cancer this dysregulated growth arises from cells in the terminal lobular unit.

Breast cancer is classified according to whether it is *in situ* and therefore has not breached the basement membrane (which does not have the capacity to metastasise) or it has breached the basement membrane (and has the capacity to metastasise) (Figure 1-5).

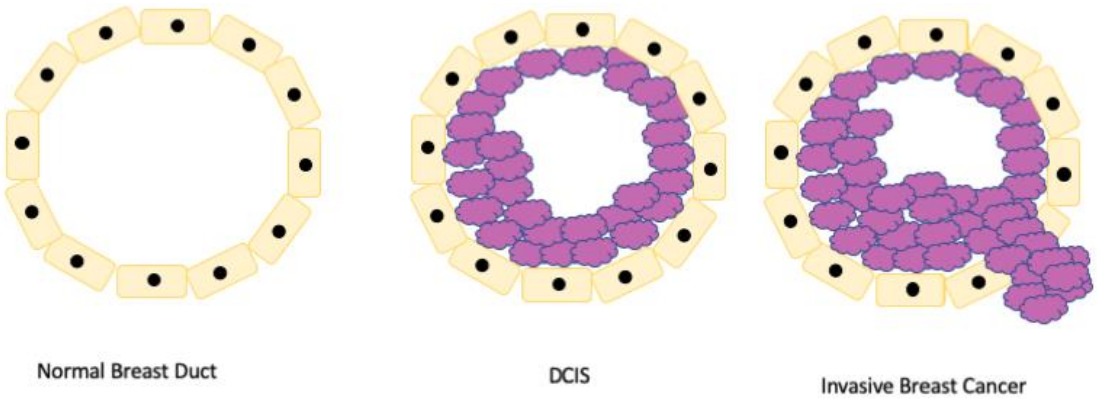


Figure 1-5 Diagram showing the difference between normal breast duct, DCIS and invasive breast cancer.

The purple cells represent abnormal 'cancerous' cells, the orange cells the basement membrane. It is only then the cancerous cells breach the basement membrane do the cells become a true cancer, with an ability for distant spread.

Non invasive (in situ) Breast cancer

This can be either Ductal or Lobular in origin (DCIS or LCIS). DCIS is the commonest form of non invasive cancer, it can be mass forming (though rare) but is associated with mammographically detectable calcifications. It therefore accounts for 3 – 4% of symptomatic cancers, and 17-25% of screen detected cancers [6, 24]. Histologically DCIS has a wide variation and is associated with pleomorphic cells with areas of calcified necrotic debris or calcified secretary material [25].

DCIS is generally treated once it is detected, and the treatment is similar for invasive breast cancer with surgical excision. The reason for treating DCIS is that it is associated with a higher risk of developing invasive cancer with a cumulative incidence of invasive breast cancer of 38% after 10 years [26], and there is a small risk of metastasis, likely due to undiagnosed small foci of invasive cancer within DCIS [25]. The natural history of pure DCIS is unknown as it has traditionally been treated; however one study looking at incompletely excised DCIS with no other treatment demonstrated that low-grade DCIS may follow an indolent course [27]. The incidence (number of new cases per year) of such lesions has increased considerably (particularly in older patients) with screening [28]. Therefore, there are concerns that breast abnormalities that may not be life limiting are being overtreated , and the LORIS trial has been established to compare outcomes of primary excision versus 10 year active surveillance in screen detected or incidental low risk pure DCIS lesions [24].

In comparison, LCIS is neither mass forming nor associated with calcifications and is nearly always an incidental finding in biopsies. LCIS has a uniform appearance with round nuclei in clusters within the lobules [25]. A diagnosis of LCIS is associated with an increased risk of breast cancer, and is managed only with regular follow up [6].

Invasive Carcinoma

The majority of invasive cancers have no special type – and are referred to as Invasive Ductal Carcinoma (IDC) and account for 85% of invasive breast cancer. Histologically this is a heterogenous group of tumours ranging from well-developed tubule formation and low grade nuclei to anaplastic cells [23]. They are commonly hard mass forming cancers, which replace normal fat tissue. There are a number of special subtypes such as lobular (which account for 10% of invasive breast cancer) and tubular, cribiform, papillary, mucinous and medullary (accounting for the rest) [6].

Breast cancers are histologically graded Grade I (best prognosis) to Grade III (poorest prognosis), based on the type of tumour, mitotic rate and degree of cellular differentiation.

Despite the wide range of histopathological characteristics which do have a prognostic significance, these histological subtypes of invasive breast cancers are diagnosed and treated in a similar may and will be considered as a single group in the rest of this thesis.

1.1.3.2 Molecular Subtypes

Even before modern medicine, it was recognised that breast cancer was a heterogenous disease as patients could present in very similar fashions, but some progressed rapidly, whereas some patients stayed alive for many years (up to 20 years) without treatment [26, 29]. An increasingly clinically relevant pathological distinction between breast cancers is receptor expression. If a cancer expresses a particular receptor to a higher degree than its surrounding normal tissue, this can act as a target for systemic treatment, able to treat local and metastatic disease.

Hormone receptors

The (o)Estrogen receptor (ER) is present in 75% of breast cancers and can be targeted in treatment. The majority of ER positive tumours also present progesterone receptors (PR), this cannot be targeted therapeutically, but has prognostic significance. A tumour that is ER and PR positive, has a better prognosis than a tumour that is ER positive, but PR negative.

Growth Factor Receptors

Growth factor receptors show increased expression and activation in breast cancer. Tyrosine kinase receptors such as the Human Epidermal Growth Factor (HER) receptors have been therapeutically targeted; HER2 being the most relevant. A tumour that is HER2 positive is a biologically poor prognostic indicator, but with modern treatments (Tranztuzamab) that target this receptor, this is no longer necessarily the case [30].

Genomic Profiling

Recent advances in genomic profiling of the molecular subtypes of breast cancer are revealing that these biochemical differences in tumours also signify underlying differences in cancer biology and can be subtyped to identify the prognostic significance between groups and an ability to predict response to therapies.

Gene expression profiling based on receptor expression has developed to show four different prognostic groups [31] , Figure 1-6 shows the difference in survival between the groups, which are:

Luminal A – High level of ER and a low proliferation rate. HER2 -ve.
Luminal B – Lower level of ER and a high proliferation rate. HER2 -ve.
HER2 +ve (or ERBB2+) – HER2+ve and either ER + or -
Basal like – ER -, PR -, HER2 -ve – so called ‘triple negative’ tumours and tend to have the poorest prognosis.

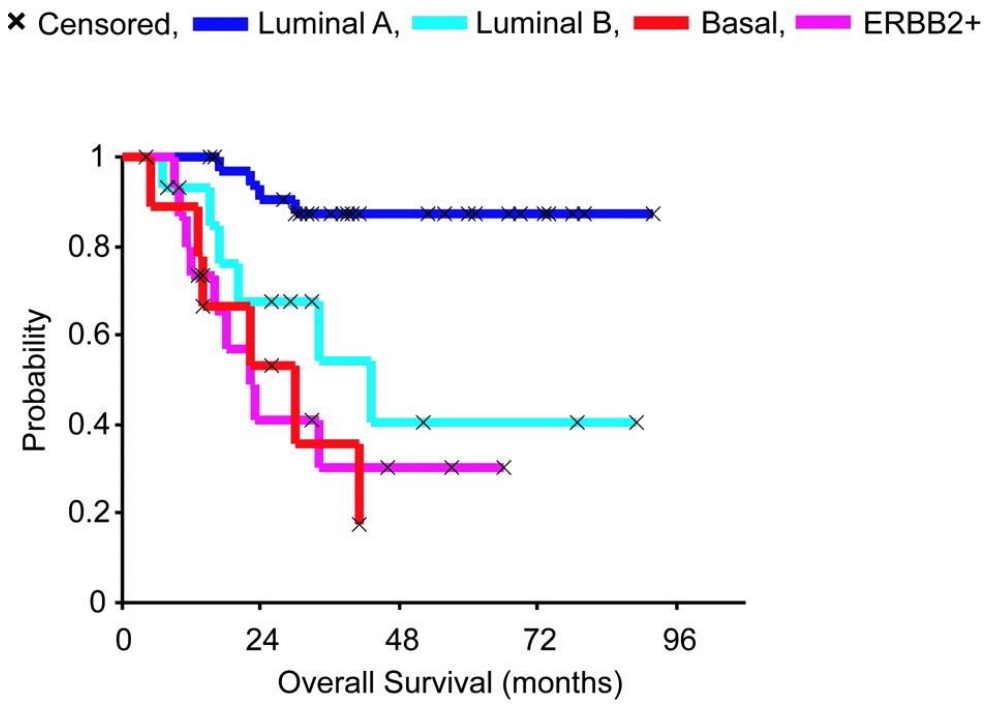


Figure 1-6. Kaplan Meier survival graph of overall survival for 72 patients with locally advanced breast disease in a Norway cohort according to molecular subtype.

Reproduced from figure from original source [32] Copyright (2003) National Academy of Sciences

1.1.3.3 Other Prognostic factors

Alongside these molecular classifications of breast cancer, there are more traditional prognostic factors which have been known about for decades:

Tumour size – the larger the tumour, the greater the risk of metastasis and a worse prognosis [33].

Lymph node spread – The number of axillary lymph nodes involved with tumour. The more lymph nodes, the worse the prognosis, with > 4 lymph nodes having a particularly poor prognosis [34]. This is a prognostic marker, rather than the simple idea that all metastatic activity occurs through the ipsilateral lymph nodes, exemplified by the fact that patients with node negative disease can still develop metastatic disease [35].

Histological grade- those with a Grade I tumour have a significantly better survival than those with Grade II or III tumours [36].

These factors can then be used to inform prognostic classification systems which are used to guide treatment decisions. The commonest used methods are the T(umour) N(ode) M(etastasis) staging, and the Nottingham Prognostic index (NPI).

The TNM classification is an evolving method of stratifying tumours. There is a ‘clinical’ stage which is based on pre- operative staging, and this can be modified with the subsequent pathological diagnosis from the excised tumour, denoted by a pre-script ‘p’. The TNM classification ascribes a number to each category, the combination of which then stratifies the disease into a Stage grouping. A modified, and simplified description of clinical TNM staging is described in Figure 1-7. A criticism of using this staging classification system is that it relies on clinical measurements, which are inaccurate [6].

T(umour) Primary Tumour	
T_{is}	Carcinoma in situ e.g. DCIS
T_1	Tumour <2 cm in greatest dimension
T_2	Tumour 2-5 cm in greatest dimension
T_3	Tumour >5cm in greatest dimension
T_4	Tumour of any size with direct extension to chest wall or skin
N(odes),Regional Lymph nodes	
N_0	No regional lymph node metastasis
N_1	Metastases to movable ipsilateral level I,II axillary lymph nodes
N_2	Metastases to ipsilateral axillary lymph nodes that are fixed or to ipsilateral internal mammary nodes
N_3	Metastasis to ipsilateral level III axillary lymph nodes, internal mammary nodes AND axillary lymph nodes, any supraclavicular lymph nodes
M(etastasis)	
M_1	No distant metastasis
M_2	Distant metastasis

Stage	**T**	**N**	**M**
0	Tis	N0	M0
I	T1	N0	M0
IIA	T1	N1	M0
	T2	N0	M0
IIB	T2	N1	M0
	T3	N0	M0
IIIA	T0-2	N2	M0
	T3	N1 or 2	M0
IIIB	T4	N0,1,2	M0
IIIC	Any T	N3	M0
IV	Any T	Any N	M1

Figure 1-7 TNM staging classification and stage grouping.

Modified from the International Union Against Cancer TNM Classification of Malignant Tumours (7th Ed) [37]

The Nottingham Prognostic Index (NPI) uses similar information, but from the pathological diagnosis only, in Equation 1-1 [38]:

$$NPI = [0.2 \times S] + N + G$$

Equation 1-1 The Nottingham Prognostic Index

S = size of index lesion (cm), N = Nodal status (0 nodes = 1; 1-3 nodes = 2, >3 nodes = 3), and G = Grade of Tumour (Grade I = 1, II = 2, III = 3)

The resultant score can give an estimated 5 year survival [39]. Both of the TNM and NPI tools are used regularly to guide management and inform prognosis. Increasingly online tools such as Adjuvant! Online and Predict! are used to determine prognosis and how different treatments would be expected to change survival rate. The clinician inputs a wide range of the patient and tumour characteristics, and the algorithm predicts the patient individualised prognostic outcome [40].

1.1.4 Diagnosis

Patients are diagnosed with Breast cancer in a secondary care setting having received a tissue diagnosis. There are two routes through which the vast majority of Breast cancer is diagnosed, and they represent different patient groups, and often different phenotypes of Breast cancer. These are symptomatic or screening.

1.1.4.1 Symptomatic

Despite screening, the majority of patients diagnosed with breast cancer are diagnosed through another route, primarily after referral to a breast unit with symptoms. Patients who present with breast symptoms and demographics that are associated with a risk of breast cancer are referred to secondary care through a fast track service for assessment (Figure 1-8) [41]. Speed of assessment is critical, as a delay in presentation and treatment of breast cancer is associated with a worse survival outcome [42]. Symptoms of Breast cancer include breast lump (the commonest), breast asymmetry, nipple retraction and nipple discharge [6]. Patient symptoms are taken in context of their other risk factors, for example the positive predictive value (PPV) of a breast lump in women aged 40 -49 being breast cancer is 4.8%, compared to a PPV of 48% in women aged >70 years [43]. Patients undergo ‘triple assessment’ during this clinic, which is discussed later.

Refer people using a suspected cancer pathway referral (for an appointment within 2 weeks) for breast cancer if they are:

- Aged 30 and over have an unexplained breast lump with or without pain or
- Aged 50 and over with any of the following symptoms in one nipple only:
  - Discharge or
  - Retraction or
  - Other changes of concern

Consider a suspected cancer pathway referral (for an appointment within 2 weeks) for breast cancer in people:

- With skin changes that suggest breast cancer or
- Aged 30 and over with an unexplained lump in the axilla

Figure 1-8 Summary box outlining the 'fast track' referral criteria from primary care to secondary care for urgent evaluation by a Breast specialist[41]

1.1.4.2 Screening

Around one quarter of the 55,000 cancers diagnosed per year in the UK are diagnosed after screening [41]. Screening invitations in the UK are currently offered to all women aged 50-70, with a phased rolling out to women aged 47-73. Breast cancer screening involves a single mammogram once every three years [44]. 1 in 23 women are called back after attending their screening mammogram for further tests, such as repeat mammogram, MRI or biopsy of a suspicious lesion [45]. Breast cancer screening reduces mortality, as cancers are diagnosed at an earlier, and more treatable stage [46]. Treating breast cancer at an earlier stage of the disease also reduces the overall cost of disease treatment, as it is less likely to require expensive adjuvant treatments [47] . Screening prevents 1 breast cancer death for every 200 women who are screened regularly [45] and in a meta-analysis of 11 studies they found that there was a reduction of 20% in the relative risk of mortality from breast cancer in screening patients compared to those not screened [48].

Despite this, screening is controversial as there are concerns over a number of issues. Over-diagnosis, where a non-life threatening breast cancer is diagnosed and treated, is a concern as this causes the morbidity of breast cancer treatment with no survival benefit. It was estimated in a Danish trial that there was 2.1 cases

of overdiagnosis per 1000 patients screened [49] and in an analysis of all nations offering screening, there was on overall rate of 1 in 3 breast cancers being over diagnosed [50]. False positive mammograms leading to unnecessary further investigation can cause distress and reduces the chance of the patient attending subsequent screening invitations [51]. Other risks include false reassurance leading to subsequent delayed presentation [52], pain and discomfort of mammograms [53], and the small radiation risk from mammograms (there are an estimated 0.07 radiation induced breast cancers per 1000 women screened) [45].

For these reasons, current uptake of Breast screening is around 70% nationally, with many regions falling below the minimum standard for uptake [54]. Some nations are re-thinking breast cancer screening, with Switzerland recently abolishing its breast screening programme citing over-diagnosis concerns as a primary reason [55].

1.1.4.3 Triple assessment

After being referred to a symptomatic clinic or re-called after screening, patients are clinically assessed through a history and examination by a specialist Breast surgeon/physician, have imaging which is either a mammogram or ultrasound or both, depending on their age [56], and if there is a discreet lesion – a core biopsy is taken for histopathological diagnosis.

Imaging

Mammogram is the primary mode of diagnostic imaging for breast cancer detection both in screening and symptomatic patients. With the patient standing, the breast is compressed between paddles and two X-ray images (cranio-caudal and mediolateral oblique) per breast are obtained [6]. In one study of 2020 patients with a subsequent pathological diagnosis of breast cancer, mammogram alone had a sensitivity of 82.9% and specificity of 91.9% for breast cancer detection, compared to a combination of mammogram, ultrasound and clinical examination (which is usual clinical practice) [57]. Factors particularly affecting diagnostic accuracy are - increased breast density, heterogenous breast density and young age [58]. For these reasons it is not routinely used in patients less than 35 years old [56].

Ultrasound is used in addition to mammograms for breast cancer diagnosis, or for those under 35 years old as the primary modality. Patients are positioned lying on a couch, and the breast is not restricted, and so it is used for guiding biopsies, or localised tumours prior to surgery. Ultrasound alone has a sensitivity of 87.6% and a specificity of 95.5% when compared to the usual practice of US, mammogram and clinical examination of the breast [57], and is particularly useful in differentiating between solid and cystic lesions [6]. The use of both mammogram and ultrasound for detection of breast cancer in a screening population results in an improved sensitivity and specificity compared to the techniques used alone [59].

Pathology

Biopsy samples are now recommended to be taken by core biopsy or vacuum assisted core biopsy as this preserves cellular architecture and increases diagnostic yield compared to traditional fine needle aspiration (FNA) [44, 60, 61]. Biopsies are taken from areas of suspicion either by clinical palpation for a palpable mass undetected on imaging, or under image guidance if detected by imaging such as for non- palpable masses and areas of suspicious heterogeneity (such as distortion or asymmetric density areas of calcification on the mammogram) [44, 62].

Once taken, biopsy tissue must be fixed in formalin as soon as possible, as a delay in formalin fixation by just an hour can reduce the sensitivity of techniques in elucidating biomarkers such as progesterone and HER2 receptor status [63]. After fixation, samples are cut and embedded in paraffin and haematoxylin and eosin stained. A histopathologist will then analyse the slide and perform any necessary immunohistochemistry to be able to provide the tumour classification, grade of tumour, receptor status [64]

Although histopathology is the current gold standard for the tissue diagnosis of breast cancer, there are issues with the technique. It is slow, the standard time for turn-around of a biopsy sample is two weeks. Due to the number of processes involved in the analysis of a specimen there are a number of points of potential human error [65, 66]. There is also a degree of inter-observer variability in aspects of histopathological analysis, in a study of 115 pathologists independently reading breast pathology slides, there was only 75% agreement between study subjects

diagnosis with the expert consensus derived diagnosis [67]. A particular issue in breast cancer is the reporting of the grade which is highly variable – in one study with six pathologists, the statistical agreement on grade diagnosis varied from moderate to substantial [68], and in a study involving five pathologists, absolute agreement was only agreed on in 83% of cases [69].

Multi Disciplinary Team Meeting (MDT)

The MDT is an important concept in cancer care in the UK, and is designed to reduce the unwarranted variation in patient care in the NHS [70]. It is a weekly meeting with a representative from all clinical members involved with breast cancer care attendance, particularly breast surgeons, radiologists, pathologists and breast cancer nurses. All patients with a core needle biopsy result should be discussed at an MDT [44]. During this meeting, all information regarding the patients' diagnosis is presented to agree on a diagnosis and suggest a management plan. Treatment is based on patient characteristics, such as co-morbidities and patient wishes, and tumour characteristics such as TNM stage, NPI and grade of tumour.

There is a paucity of evidence demonstrating the efficacy of the MDT [71]. And a variety of evidence to suggest that not all patients suitable for MDT are discussed in the meeting [72], there is a large variation in their use between countries [73] and an inability to reach consensus decision in 27-52% of cases in one systematic review [74]. Despite these criticisms, an effective MDT improves patient outcomes. A systematic review of 12 studies demonstrated an improved patient survival after the introduction of a cancer MDT [75], and the mortality rate was found to be 11% lower in an area of Scotland that introduced an MDT compared to an area with similar survival statistics prior to the intervention that did not introduce an MDT over the same time period [76].

1.1.5 Management of early breast cancer

Surgery remains the first line treatment in the majority of early and locally advanced breast cancer. Recent advancements in adjuvant and neo – adjuvant therapies have significantly changed breast cancer management and outcomes.

1.1.5.1 Adjuvant therapy

The recent decrease in mortality from Breast cancer [14], despite a rising incidence, is likely due to the widespread adoption of adjuvant systemic therapies. The aim of adjuvant therapy is to improve disease free survival and reduce local recurrence rate, whilst taking into account the (sometime significant) side effects. Breast cancer is a heterogenous disease, and the precise management decisions are tailored to the patient and the biology of the tumour, which is decided at MDT. A comprehensive review of all possible adjuvant therapies and ongoing clinical trials is beyond the scope of this thesis, what follows is a highlighting of the common treatment types.

Endocrine therapy

Around 75% of breast cancers are ER+ve [77] and modulation of ER activation in these cancers is essential. This includes stopping any oestrogen containing contraception or Hormone replacement therapy, and consideration of ovarian suppression with GnRH agonist in high risk pre-menopausal women [78].

Tamoxifen – is a competitive oestrogen receptor antagonist, blocking estradiol from binding and activating the estrogen receptor. It reduces the risk of death by around one-third during 5 years of treatment, with some benefit up to 10 years [79]. NICE recommends offering tamoxifen in ER positive invasive breast cancer for men, premenopausal women and postmenopausal women with low risk of disease recurrence [56].

Aromatase Inhibitors (AI)– the mechanism of action is by inhibition of the enzyme aromatase that converts androgens into oestrogens. In the ATAC trial, looking at an AI and tamoxifen alone or in combination demonstrated that AI alone was superior in terms of disease free survival [80], and another trial demonstrated a possible survival advantage for those taking AI alone [81]. AI is contraindicated in premenopausal women and have significant side effects of sexual dysfunction, and NICE recommends AI in postmenopausal women with ER positive disease at medium or high risk of recurrence [56].

Chemotherapy

The recommended chemotherapy regime in the treatment of breast cancer is a combination containing a taxane (e.g. docetaxel) and anthracycline (e.g. doxorubicin)[56]. It is an effective treatment in early breast cancer and can reduce

the annual risk of recurrence by almost 25% and reduce breast cancer mortality by a third [82]. The decision to initiate chemotherapy is based on the benefits of survival versus the significant side effects, and is generally offered in higher risk patients. The gain from chemotherapy is higher for younger women than in older women. A number of factors is taken into account by the MDT in deciding whether to offer chemotherapy, this includes age, axillary nodal status, ER status, and DNA micro analysis in the form of Oncotype DX [6].

Biological Therapy

Trastuzamab is a recombinant humanised monoclocal antibody that binds to the HER2 receptor and antagonises its activation. It is recommended in HER2 positive tumours in the adjuvant setting [56], and when given concurrently with chemotherapy leads to a decrease in recurrence and death [83] in these tumours that classically had a poor prognosis.

Radiotherapy

Radiotherapy is a local treatment (compared to the systemic treatments of other adjuvant therapies) and is a part of breast conserving therapy and is always given to the remaining breast tissue after breast conserving surgery [84]. Depending on the disease characteristics and staging, radiotherapy is given for locally advanced tumours and can be given to the tumour bed, axilla, supraclavicular fossa and inframammary chain [56]. Radiotherapy significantly reduces locoregional recurrence and improves survival in locally advanced disease [85].

1.1.5.2 Surgery

The aim of an oncological resection is to remove the minimal amount of tissue necessary to reduce trauma and collateral structure damage, whilst excising the entirety of the diseased tissue [86].

Mastectomy

A mastectomy is the removal of all breast tissue of a single breast, and is the traditional operation for the treatment of breast cancer. Originally a Halstead mastectomy included removal of the underlying pectoral muscle and all axillary lymph nodes as well, but with a recognition that this did not always improve survival and caused significant morbidity, the technique evolved to what is now a

modified ‘simple’ mastectomy. It accounts for 30 – 40 % of breast cancer operations. Local recurrence can occur in any residual breast cells which may remain, or associated with the scar.

Breast Conserving Therapy

With the increasing realisation that breast cancer was a systemic disease, the improvement of adjuvant therapies, and the detection of smaller cancers through mammography and screening, less invasive surgical treatments were developed. Breast conserving therapy consists of Breast Conserving Surgery (wide local excision of the tumour) – the aim being to remove all cancerous tissue with a margin of normal surrounding tissue, followed by whole breast radiotherapy. The surgeon makes an incision either over the cancer, or in a more cosmetically sensitive place on the breast and excises the breast cancer in a lump, with a rim of healthy tissue (roughly aiming for 1 cm of normal feeling tissue around the cancer), with the aim to excise a specimen with pathologically clear margins (Figure 1-9).

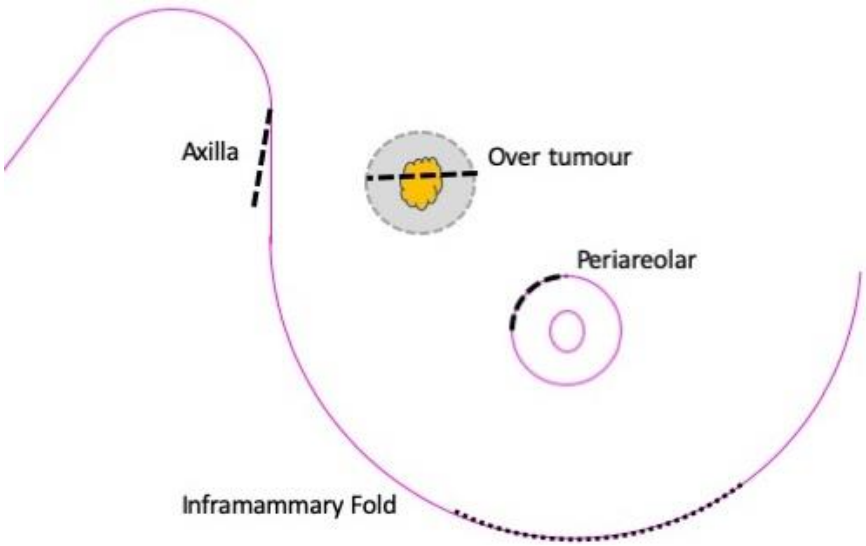


Figure 1-9 Diagram demonstrating options for incisions for Breast conserving surgery.
Yellow area denotes location of tumour, grey shaded area is the area of tissue resected by surgeon (to include tumour and rim of healthy tissue). Black dotted lines are potential incisions made for tumour resection.

The original randomised trial demonstrating the safety of breast conserving therapy was by Fisher *et al.* where 1843 women were assigned to one of three

groups – total mastectomy, segmental mastectomy, and segmental mastectomy with breast irradiation [84], this study at the 5 year follow up actually found improved disease free survival and overall survival with segmental mastectomy and irradiation compared to total mastectomy. There are differences between the study conditions and current practice which could limit these conclusions, such as greater breast tissue volume excision than current lumpectomy, higher radiation doses, and that all patients received systemic chemotherapy. However, the findings have been repeatedly confirmed with more current practices, and a meta-analysis of nine trials comparing mastectomy with breast conserving therapy demonstrated no significant difference in 10 year mortality [87], and review of six randomised controlled trials showed no difference in rates of local control [88], and so mastectomy and breast conserving therapy are viewed as equivalent in providing oncological control of disease.

Indications for breast conserving therapy are T1, T2 (tumour < 4 cm), NO, N1, M0 tumours, T2 tumour > 4 cm in large breast, unifocal lesions. Contraindications are T4, N2 or M1 lesions, multifocal disease, large tumours in small breasts and women with a strong family history or proven BRCA genetic mutations [6]. The advantages of breast conserving therapy compared to mastectomy are related to lower psychological morbidity with less anxiety, depression, and an improved body image, sexuality and self-esteem [89, 90].

A significant issue with BCS is that incomplete excision and subsequent positive pathological margins lead to a second operation for re-excision of margins in around 17% of patients in the UK [91] to achieve clear resection margins, which is necessary to reduce recurrence rate [92]. The current gold standard for margin analysis and the rationale for need for intraoperative margin analysis is explored in greater detail in section 1.3 of this thesis.

Axillary Surgery

The management of the axilla in breast cancer continues to be controversial. Although having metastatic carcinoma in the ipsilateral axillary lymph nodes has undoubted prognostic significance [34], the removal or treatment of these lymph nodes with radiotherapy may not carry a survival benefit [93, 94]. Therefore, axillary surgery is primarily for staging, and in those with a positive axilla, local control.

Sentinel lymph node biopsy – this is performed in all patients with invasive breast cancer with a pre-operative diagnosis of a node negative axilla [95]. Radioactive and blue dye is injected into the breast to identify the first node(s) that the breast drains, and these lymph nodes are excised for staging purposes [96]. If the excised lymph nodes do not contain cancer, no further axillary treatment is required. If the excised lymph node(s) do(es) contain cancer the patient may be offered axillary lymph node clearance or axillary radiotherapy [94] (this is an area of active clinical trials [97]).

Axillary Lymph Node Clearance – in patients who are pre-operatively identified as having a node positive axilla with US imaging and pathological confirmation from a core biopsy, the treatment is currently for an axillary lymph node clearance [56].

1.1.6 Summary

Breast cancer is a potentially life-limiting disease, and due to the high, and increasing, incidence the disease and its treatment is a significant burden to populations and healthcare systems globally. Although there have been advancements in treatments and survival has improved, there remains a clinical need for improvements in delivering cancer care. Breast Conserving surgery is the commonest operation performed for breast cancer, and the issue of positive resection margins requiring re-operation is a significant source of anxiety to patients, and an expensive burden for healthcare systems. If the resection margins of a lumpectomy specimen could be analysed at the time of the index operation -intraoperative margin analysis (IMA)- this could reduce the need for repeat operations, and improve patient cancer care. The Association of Breast Surgeons gap analysis working group recently produced a report highlighting that research into technologies that could provide IMA were a priority for surgical research [98]. Raman Spectroscopy is a method of vibrational spectroscopy that can assess the disease specific molecular compositional changes in the cells and tissues, and the technique holds great promise in being able to perform IMA.

1.2 Background of Raman Spectroscopy

1.2.1 The theory of Raman scattering

Raman scattering is the inelastic scattering of light. When light is directed on a molecule, light may either pass through it without any interaction or may interact with the molecule – where two processes can occur. The molecule can either absorb energy from the incident photon and it is transferred to a higher energy state (which is measured in absorption spectroscopy) or it can scatter the incident photon, which is measured in Raman spectroscopy.

Light as a form of radiation may be considered both in terms of a wave (with a wavelength) and as a particle (a photon) with a discreet amount of energy. The amount of energy the light has is related to the wave properties. The energy of a photon (E) is related to wave frequency (*v*) by Equation 1-2:

$$E = hv$$

Equation 1-2 Where h = Plancks constant (6.62607004X 10 $^{-34}$m^{2}kg/s)

The wave frequency (*v)* is inversely related to the wavelength (λ) by Equation 1-3:

$$\lambda = {}^{c}/_{v}$$

Equation 1-3 Where *c* = the velocity of light (299,792,458 m/s^2)

Therefore, wave frequency is linearly related with energy (Equation 1-4):

$$\nu = {}^{\Delta E}/_{h}$$

Equation 1-4

and wavelength is inversely related to energy, thus the shortest wavelength confers the greatest energy as shown in Figure 1-10:

Radiation Type	Gamma Ray	X Ray	Ultraviolet	Visible	Infrared	Microwave	Radio
Wavelength (m)	10^{-12}	10^{-10}	10^{-8}	0.5^{-6}	10^{-5}	10^{-2}	10^{3}
Energy per photon	1.24 MeV	1.24 keV	12.4 eV	1.24 eV	0.124 eV	1.24 μeV	1.24 peV

Figure 1-10 The Electromagnetic spectrum on the wavelength scale in metres with associated Energy per photon in electron volts (eV)

The Raman effect was first described experimentally by Raman in 1928 by focusing sunlight through a telescope onto samples of purified liquid or dust-free vapour, and observing scattered radiation with a different frequency to the incident light [99]. Nowadays a single wavelength of light is used to provide a coherent wavelength as an incident beam to irradiate the sample, and the energy changes that can be detected are more subtle, but the principles remain the same.

At room temperature the majority of molecules will be in a low energy, ground state. When an incident light source interacts with a molecule and the incident photon interacts with the molecule it forms a short-lived state of excitement due to its polarization of electrons and is promoted to a 'virtual state', which is unstable, and a photon is re-radiated. Elastic scattering accounts for the majority of scattering and occurs when there is only electron cloud distortion involved in the scattering process. It is common, and due to the light weight of electrons, results in very small frequency change to scattering photons and the molecule returns to its ground state and so the scattered photon is of equal energy to that of the incident photon. This is called Rayleigh scattering in molecules. Inelastic Raman scattering occurs when the energy transfer is significant enough to cause nuclear motion. There is energy exchange between incident photon and molecule, or molecule to scattered photon. Once the molecule has temporarily been promoted to a higher energy excited vibrational state it returns back to an energy state that is different to its original state. In this process there is a difference of one vibrational unit between incident photon and scattered photon, and this is Raman scattering. In Raman scattering there is Stokes and anti-Stokes scattering. During Stokes scattering the molecule absorbs energy and is promoted to a higher energy excited vibrational state. Some molecules may already be in an excited energy state due to thermal energy, and when these molecules return to ground state this will result in a scattered photon – this is anti-Stokes scattering. These energy exchanges are depicted in Figure 1-11. For most molecules, at room temperature, the number of molecules at the elevated excited energy state would be very small, thus Stokes scattering will be prevalent.[100]. Although highly specific, Raman scattering is considered weak, with only one Raman scattered photon in every 10^6- 10^8 of total scattered photons.

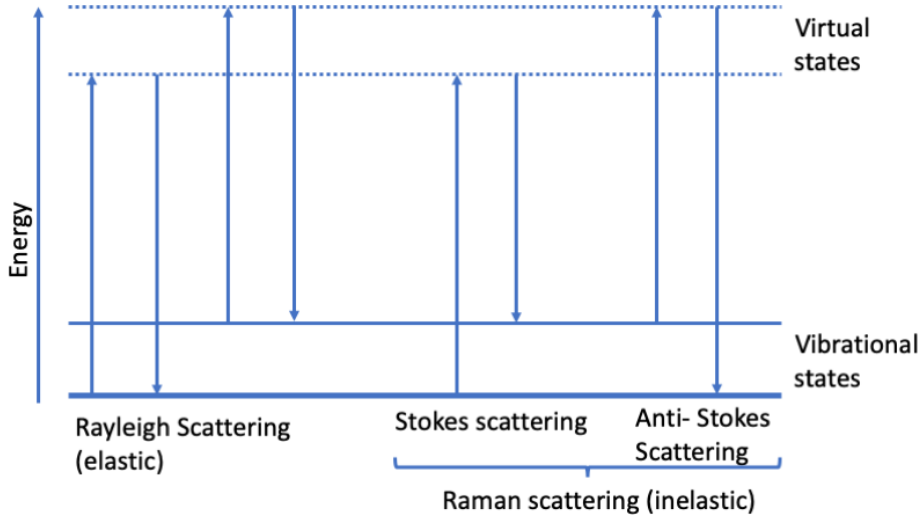


Figure 1-11. A Jablonski energy diagram of quantum energy transitions for Rayleigh and Raman scattering.

In Rayleigh scattering the energy level returns to its original state and there is no change in energy. In Raman scattering the energy level returns to a different state- either gaining (anti-Stokes scattering) or losing (Stokes scattering) energy and so the scattered light has a different wavelength to that of the incident light, which can be measured as the Raman shift from the Rayleigh line (which is equal to the incident light wavelength).

In order for a molecule to scatter light it requires polarizability, which is the ability of an electric field to induce a dipole moment in a molecule. The induced dipole moment (μ) is equal to the molecule polarizability (α) and the incident electric field (E) ; Equation 1-5 [101]:

$$\mu = \alpha E$$

Equation 1-5

A change in molecule polarizability is therefore required to change the amplitude modulation of the dipole moment that is necessary for Stokes and anti-Stokes emission. The change in molecular polarizability occurs through molecular vibration and is a result of the alteration in the electron cloud about the molecule. The molecular vibration is described by translation of the molecule in space, and rotational movement. The degree of symmetry defines how Raman active a molecule will be, with symmetrical and nonpolar groups being the most active [101]. Therefore the molecule, and the molecule number (N), the polarizability of the molecule (α), the incident laser intensity (I_0), frequency of the laser (v), and the vibrational amplitude (Q) all determine the intensity of the Raman scattered radiation (I_R), the relationship is shown in Equation 1-6:

$$I_R \propto v^4 I_0 N \left(\frac{\vartheta\alpha}{\vartheta Q}\right)^2$$

Equation 1-6

From this it can be concluded that by increasing the frequency of the laser, or using a shorter wavelength can increase the Raman intensity. It is also demonstrated that a change of molecular polarizability is essential for a molecule to be Raman active. [101]

The degree or type of scattering, and the gain or loss of energy, is dependent on the degree of movement a molecule has which is a product of its atoms and bonds, resulting in each molecule producing a unique Raman spectrum, or 'fingerprint'[100] . In a steady electrical state, the energy of a molecule is divided into different 'degrees of freedom'. Three degrees of freedom describe the translation of the molecule in space, and three describe rotational movement. Molecules thus undergo a number of bond vibrations, symmetrical (when it is symmetric about the centre) vibrations causing the greatest electron cloud distortion and a greater Raman scattering. Bond vibrations include symmetric stretching, asymmetric, stretching and are shown in Figure 1-12.

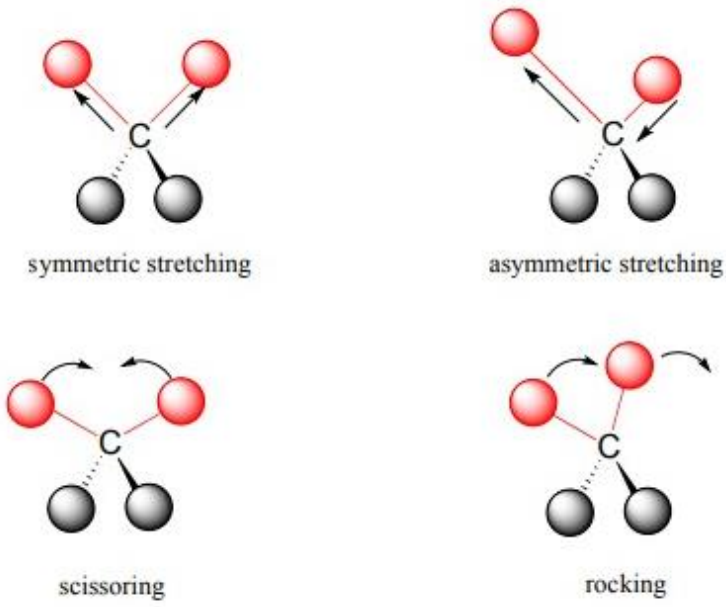


Figure 1-12 Schematic diagram demonstrating the different modes of bond vibrations that may occur within a molecule.

Reproduced from [102] under a creative commons 4.0 licence https://creativecommons.org/licenses/by-nc-sa/4.0/

The molecular bond can be mathematically considered to be a spring between two masses, and so relationship between the weight of the atom, and the strength of the bond dictate the frequency of the vibration, is explained in Hookes law in Equation 1-7:

$$\upsilon = \frac{1}{2\pi c}\sqrt{\frac{K}{\mu}}$$

Equation 1-7

Where *c* – the velocity of light, *K* is the force constant of the bond between atoms A and B (which is measure of bond strength), and μ is the reduced mass of atoms A and B of masses M_A and M_B which are related by Equation 1-8;

$$\mu = \frac{M_A M_B}{M_A + M_B}$$

Equation 1-8

And it can be seen that the lighter the atom, and the stronger the bond, the higher the frequency. The product of these features, along with the different types of vibration, will therefore give different energies/wavelengths and different intensities for each molecule.

The unique vibration of an atom or bond therefore gives a peak of increased intensity of scattered light at a particular wavelength on the spectrum. As the wavelength of a particular peak will differ according to the wavelength of the incident light, the Raman shift (the energy difference between incident light and the scattered (detected) light) is usually used to describe peaks to allow comparison of data between different incident lights. As molecules are a result of a collection of bonds formed between atoms which interact and influence their vibrations, more commonly there are ‘group vibrations’ and spectral peaks are assigned to groups – such as the ‘aromatic C-H’ or ‘O-H’ regions. As molecules are a consequence of the unique make up of a combination of these vibrational groups each molecule will have a unique Raman spectrum. The resulting peaks of differing energies and intensity can be assigned to known vibrational groups, and the molecular composition of a molecule can be deduced. Reference libraries now exist of known energies and frequencies assigned to groups to allow spectroscopists to interpret their spectra [103]. From a biological perspective, the majority of molecular information is gained within the ‘fingerprint’ region, which is between 400-1800 cm^{-1} Raman shift. In this region there is well documented

assignment of nucleic acids, proteins and lipids – which form the bulk of the information that can be gained from Raman spectroscopy. There is an ability to identify other biomolecules such as hydroxyapatite, porphyrins and carbohydrates [104].

The High wavenumber (HWN) region corresponds to Raman shift 2700-3800 cm^{-1}, which is highly sensitive to lipids, proteins and water [105] but is often seen as less specific than the fingerprint region. However, it has a good ability to differentiate between a number of different cellular components [106] as demonstrated in Figure 1-13. And the diagnostic ability of the HWN region has been shown to be equal to the fingerprint region in clinical samples [107]. Between the Fingerprint and HWN regions there is little useful spectral information.

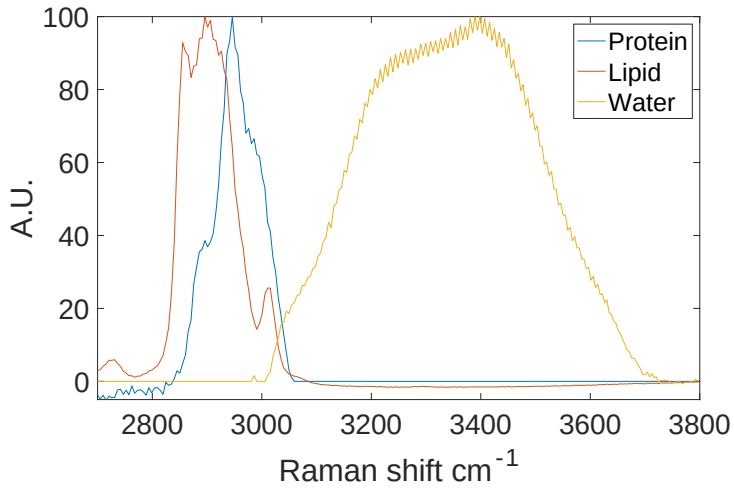


Figure 1-13 Representative Raman spectra demonstrating the cellular molecular components that can be measured using high wavenumber Raman Spectroscopy.

Raman spectra taken of Bovine gelatine (protein), soya bean oil (lipid) and distilled water (water) with 785nm laser excitation

1.2.2 Experimental measurement of Raman scattering

There are a number of systems used to measure Raman scattering, and there are a number of techniques that are discussed later in this thesis, each of which can be used to provide different information about the molecule and biological tissue probed. However, all systems have the same basic simple components, which are demonstrated in Figure 1-14.

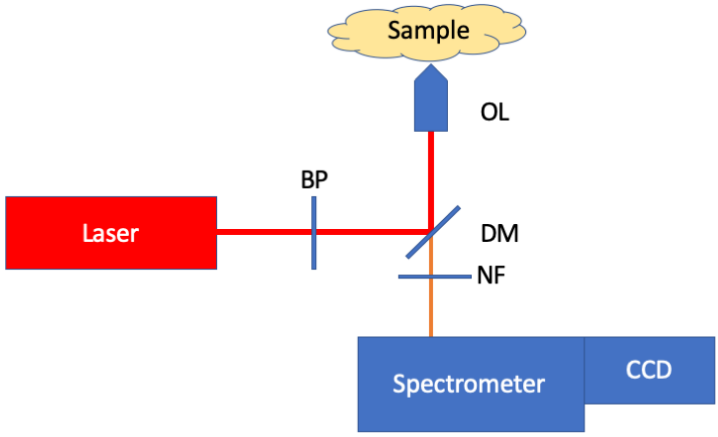


Figure 1-14 Diagram demonstrating a typical set up for a system to perform Raman spectroscopy.
OL= Objective lens, BP = bandpass filter, NF = notch filter, DM = dichroic mirror

Laser

The excitation light source in Raman spectroscopy is usually a laser – as this provides a single narrow band monochromatic light source to probe the molecule. Lasers are the most suitable light source because a large amount of light is needed to elicit Raman scattering due to its weak effect, and the unique specificity of a Raman spectra is dependent on a single wavelength and a broader light source would give a broader, overlapping, spectra that would be difficult to interpret [104]. The commonest type of laser used in Raman spectroscopy of biological material are diode lasers that converts electric energy into light energy. From Equation 1-6 it could be seen that by decreasing the laser wavelength would increase Raman intensity, however there is a balance when dealing with biological samples, as shorter wavelengths can heat specimens and cause molecular change or damage. It is also at these wavelengths that the more readily detected effect of auto-fluorescence can become an issue.

The Near-Infrared spectrum is commonly utilised in Raman spectroscopy, particularly of biological tissue samples, as it achieves the optimum balance between intensity, tissue penetration and confounding auto-fluorescence [108]. It also appears to be the safest wavelength to use. One study looking at immortalised cell lines demonstrated rapid cell death after short exposure times (within 5 minutes) to wavelengths 488 and 514nm, but only very mild morphological changes and no cell death after exposure to 785nm wavelength for 40 minutes [109]. Another study used real-time analysis of cardiomyocytes under

a 785nm laser to demonstrate there was no effect on the beating frequency suggesting minimal effect on the cells function [110].

Optical filters

A variety of optical filters are necessary to ensure a homogenous light is transmitted and the collection device receives only scattered light. Band-pass filters effectively clean up the laser light and are used to allow only the transmission of the excitation light to the sample, ensuring monochromatic light and optimal spectral resolution. The Dichroic mirror reflects the excitation wavelength and longer wavelengths towards the objective lens – which focuses the incident light on the sample to be interrogated. The scattered light then comes back through the objective lens, the dichroic mirror allows all scattered and red-shifted light through to the collection device, while reflecting light at the incident wavelength and longer away. The notch filter provides a further filter to block light at the incident light wavelength through to the spectrometer.

Detection

Gratings can be used to aid in disperse the light into its individual wavelengths. Dispersion gratings separate the collected scattered Raman light at different wavelength onto the detector. The more grating lines per mm, the higher the dispersion and therefore the spectral resolution. The scattered light is then detected by pixels on a camera chip. Each pixel (or group of pixels) in the chip corresponds to a separate wavelength, and the amount of charge the pixel acquires during an acquisition is directly proportional to the number of photons detected. At the end of the acquisition the resultant charge collected by the pixel is converted to voltage and this is transferred to a computer to produce a signal for that wavelength, and the chip as a whole provides a read out for the entire spectral range. The commonest camera used are thermoelectrically cooled Charge – Coupled Devices (CCD). The CCD is a silicon chip made up of a 2D array of detectors. The sensitivity of a camera is dependent on the photon wavelength and the silicon absorption coefficient. This is described as the quantum efficiency of a camera and is defined as the number of photoelectrons generated per Raman photon [111]. The sensitivity of the CCD is dependent on incident wavelength and decreases above 1000nm.

If a silicon CCD is not appropriate for the spectral range being investigated there are alternatives. Indium gallium arsenide (InGaAs) cameras are typically back illuminated with a two dimensional photodiode array that, due to its lower bandgap properties, have an improved quantum efficiency over 900nm, and are generally used for investigating in the wavelength around 900 – 1700nm[112]. Alternatives to silicon CCDs also include complementary metal oxide semiconductor (CMOS) cameras, which are generally cheaper than CCD cameras, and can provide a higher frame rate.

1.2.3 Summary

The basic principles of Raman spectroscopy have been discussed. It is a versatile technique that has a wide range of applications in biomedical science. There are a variety of methods of Raman spectroscopy, which can be tailored to the specific problem posed; these are discussed in the following section. A potential application of Raman Spectroscopy is to the intraoperative assessment of pathological margins in surgically excised specimens – the next part of this thesis will give a detailed review of recent developments in this field of research.

1.3 Raman Spectroscopy for rapid intra-operative margin analysis of surgically excised tumour specimens

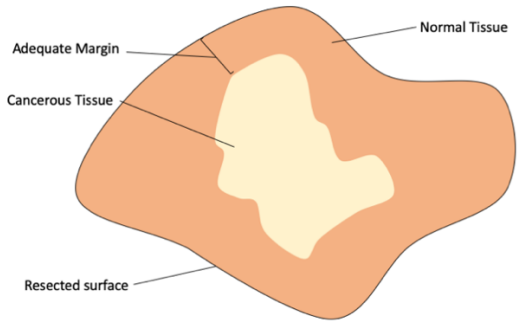
(edited version from article published in *Analyst*)

Breast Conserving Therapy with wide local excision of breast tumour is the commonest form of operation performed in the treatment of primary breast cancer, as detailed in section 1.1.5 of this Background chapter. The high rate of 'positive' pathological margins requiring a further operation for re-excision of margins in around 17% of patients represents a significant burden to the patient and to the healthcare system. Raman spectroscopy may be the ideal intraoperative tool to provide intraoperative margin analysis. This section of the background evaluates the ability of Raman spectroscopy to provide IMA, setting out the clinical need for IMA in breast cancer, and a number of other solid tumour pathologies, a current critical review of previous work done, and an evaluation of the ability of RS to provide IMA. The literature review is a modified version of the article that was published in the journal Analyst (Reproduced from Reference [113] with permission from the Royal Society of Chemistry). The introduction and section on Raman Spectroscopy background have been altered to avoid repetition within the thesis, while the rest of the text is unchanged from the published version.

1.3.1 Introduction

Surgical excision of the primary tumour remains the primary treatment for many solid organ tumours [86]. The aim of cancer surgery is to remove the smallest amount of tissue necessary to minimise tissue trauma and collateral structural damage, whilst excising the entirety of the diseased tissue [86]. This requires the affected tissue to be excised with a rim of normal tissue with an adequate 'margin'. The amount of margin, or distance from the cancerous tissue to the edge of the specimen, required to be termed 'clear' is different for each pathology. If there is cancerous tissue within the defined distance from the resected surface, it is a 'positive' margin and an inadequate resection, which increases the risk of recurrence [114-116] (Figure 1-15).

A



B

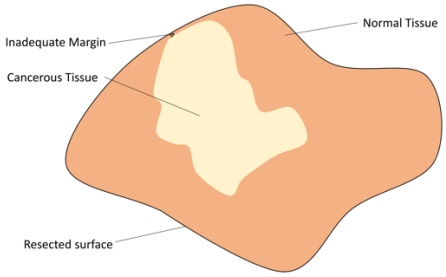


Figure 1-15 A graphic to illustrate the concept of tumour margins.

A. A surgically excised specimen with an adequate margin along the resected surface; the cancerous tissue is in the centre of the surgically excised specimen, with a rim of normal tissue surrounding it. The distance of what defines an 'adequate' margin varies between pathologies. B. A surgically excised specimen with an inadequate or 'positive' margin; the cancerous tissue is at the edge of the specimen. This could mean there is further cancerous tissue in the patient that has not been excised.

A positive margin not only affects prognosis, but also future management, meaning the need for further operations or adjuvant therapies with significant side effects. The cost of a positive surgical margin to the patient, in terms of increased treatment burden, further anxiety and additional side effects is difficult to quantify but the effect on the financial resources of the healthcare provider is undoubtedly significant [117, 118]. A method to assess the margins of the excised specimen intra-operatively to allow further tissue to be taken at the time of the initial procedure, if possible or necessary would be efficient, could reduce the risk of residual cancer at the end of the operation and improve patient care.

1.3.2 Current methods of margin analysis

The 'gold standard' method for analysis of resection margins of surgically excised specimens is currently histopathology. Histopathology analysis of prepared, *ex vivo*, tissue is conducted with light microscopy by trained physicians and is able to provide a detailed analysis of the excised specimen and the biochemical characteristics of the tumour, which contributes to clinical management decisions. However, the 'gold standard' of histopathology is prone to errors although it is likely underreported; one study found an error rate of up to 11% in cancer diagnosis [119], there is variability between pathologists in the reporting of the tumour grade [68, 69], and even variability in the final diagnosis [67]. This is even

more apparent when diagnosing early or pre-cancers, where histopathology can have of the order of nearly 50% inter-observer agreement [67, 120].

Histopathology analysis of surgical margins may also be subject to errors. Even small specimens of around 2 cm across would require an impractical number of histological sections if the entire resected surface was to be examined for adequate margins, and so margin assessment should be seen as a representative sampling procedure [121]. The process also requires a number of steps which introduce sampling error - orientation by the surgeon, fixation to preserve the specimen, labelling the specimen, transportation to the pathology department, re-orientation by the pathologist, slicing the specimen, dehydration, embedding, sectioning, staining, representative sampling and then subjective assessment by histopathologists [65, 66, 122]. In addition to this, the lack of real-time reporting delays treatment decision making meaning histopathology is an imperfect technique for the reporting of margins. This precludes any removal of further, possibly cancerous, tissue without the need for a second, often more difficult, operation. In many pathologies, the risk of a second operation outweighs the benefits of the risk of recurrence, and so in cases of a positive resection margin, potentially cancerous tissue remains [123, 124]. It is possible that the concept of requiring a defined margin for surgically excised specimens is a result of the current inability to check the entire surface margin, and that if a technique could accurately assess a specimen, smaller, or no, margin distances would be required.

1.3.3 Methods of Intra-operative Margin Analysis (IMA)

Recognising these limitations of the traditional model, there is a large body of research investigating methods of IMA. Current and prospective methods of IMA are generally based on *ex vivo* analysis of the excised specimen, as it is the most practical way of avoiding surgical artefacts such as blood contamination and the space constraints of *in vivo* analysis.The number of methods being investigated is vast, what follows is an overview of more established techniques that have had routine clinical application, though not necessarily widespread adoption.

A commonly used method in cancer surgery is frozen section analysis. The specimen is transferred to the pathology department, frozen and sections of

interest taken for analysis by a pathologist. It is a technically difficult procedure, requiring a histopathologist to be available on demand, a turnaround time that can impede surgical workflow, the processing can damage specimens which require further histopathological examination and it is expensive [125]. Frozen section is most successfully used in Moh's micrographic surgery in the treatment of Basal Cell Carcinoma (BCC) of the skin, where the entire resected surface is examined in horizontal sections intra-operatively and the surgeon continues to excise tissue until the margin is clear. Recurrence rates are as low as 1-3% even in recurrent and complex disease [126]. However due to cost and time issues it is reserved for cosmetically sensitive areas and recurrent BCC. Frozen section is used widely in head and neck cancers [127] however there is a significant number of false negatives [127] and it is unreliable for eradicating positive final margins [128]. The technique has also been used for IMA in breast cancer excision (where there is reported sensitivity 70-90% and specificity 80-90%[129, 130]) and radical prostatectomy (with a poor sensitivity of 42% [131]), but has not been widely adopted due to clinical and cost-effectiveness concerns [132, 133].

Intraoperative imprint cytology has shown promise in Breast surgery. A slide is pressed onto the lumpectomy resected surface and analysed by a pathologist for malignant cells, which can be reported within the time frame of an operation. Issues identified with the technique are that slide preparation can affect the outcome, it is less accurate in lobular carcinoma [134] and in tissue that has been subject to previous radiotherapy [135], and it reports only on the resected surface, not the entire margin depth. A meta-analysis of eleven studies revealed a pooled sensitivity of 91% and specificity 95% [136], however, in clinical trials the need for delayed re-excision remained disappointingly high [134] suggesting this may not translate into improved clinical practice.

Intra-operative imaging to assess margins has been used in various pathologies. In breast surgery the most common adjunct in the UK to analyse margins is an intraoperative specimen X – ray to determine how close radio-opaque lesions are to the edge of the sample. It is used by 96% of UK units [91], as it is readily available and requires no additional equipment. However, in a recent meta-analysis of nine studies to have a pooled sensitivity of 53% and specificity of 84% [136].

Intraoperative ultrasound has also been used to guide lumpectomy in breast conserving surgery, but margin assessment was associated with a high false positive and false negative rate [137] and poor sensitivity [138]. In brain lesions, intraoperative US has been seen as useful to guide excision, however there is a high degree of user variability and has been associated with a high rate of false positives possibly due to blood artefact which is also hyperechoic [139].

Magnetic Resonance Imaging (MRI) is used extensively in the planning of brain cancer surgery and intra-operatively to guide excision. In a randomised controlled trial, with a small population of 58 patients, the intra-operative use of MRI was associated with complete tumour resection [140]. However, an issue is 'brain-shift' where loss of cerebrospinal fluid and oedema changes the anatomy and so reduces the accuracy of neuronavigation [141] and a Cochrane review concluded that further studies into efficacy were needed [142].

The sheer number of techniques proposed for IMA demonstrates that no single method has proved to be effective enough to be accepted into routine clinical practice. To address this gap in surgical care, novel optical techniques show promise as they can provide sensitive and specific biochemical information at a molecular level in a rapid and non- destructive manner. A number of microscopy techniques, such as Microscopy with UV surface excitation (MUSE), and light sheet microscopy, show the ability to differentiate between cancerous and non-cancerous tissue but studies so far have been limited in sample size and to physically small samples due to speed of analysis, which limits the conclusions as to clinical relevance [143] [144]. Other optical techniques have struggled to succesfully translate promising laboratory work into the clinical environment, such as optical coherence tomography (OCT) which was found to have reduced effectiveness when it encountered cauterised tissue and blood [145], and a clinical trial in breast specimens demonstrated it only identified 63% of those with a positive margin [146]. Diffuse optical spectroscopy has the potential to provide excellent sensitivity and specificity in cancer diagnosis[147], but so far this has not been realised in the analysis of margins[148]. Raman Spectroscopy (RS) is a technique of vibrational spectroscopy that has gained particular momentum as it can provide detailed biochemical information with excellent accuracy, within clinically relevant times and has been succesfully used in the surgical

environment in human tissue in a range of pathologies. RS has the potential to change the paradigm of oncological surgery and provide IMA; an idealised surgical workflow of this is suggested in (see Figure 1-16).

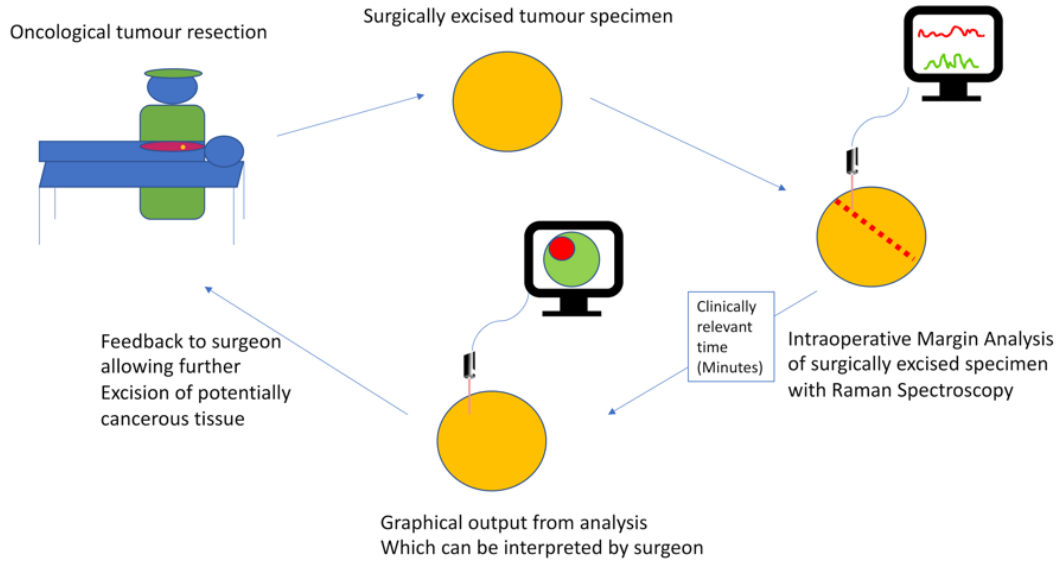


Figure 1-16 A graphical representation of the ideal paradigm for the surgical workflow of intraoperative margin analysis (IMA) by Raman Spectroscopy.

This would allow the surgeon to remove all cancerous tissue at the initial operation, thus improving patient outcome.

1.3.4 Raman spectroscopy

A detailed review of the basic principles of RS are outlined in the previous section 1.2 of the thesis. The molecular specificity of the Raman spectrum is holding great promise in medical diagnostics [149], and a variety of RS techniques have been developed in response to different needs. The following is a review of the different RS techniques that have been used in biological applications that are relevant to IMA.

Raman micro-spectroscopy

This technique combines Raman spectroscopes with optical microscopes, allowing for analysis of sub millimetre specimens, such as histological slides [150]. The tissue is scanned using raster scanning (point by point) which is time consuming and so processing a sample can potentially take hours. Line scanning (changing the size of the incident beam to a line across the sample), and multi-focal Raman microspectroscopy (dividing the incident laser into several beams to measure multiple Raman spectra simultaneously) can reduce the analysis time

to more clinically relevant times [151-153]. Wide-field global imaging techniques, where the whole sample is illuminated and all spectra associated with a particular wavelength are collected can also decrease analysis times, but flat field illumination can be difficult and this reduces the laser power per pixel, unless a very high power laser is utilised [154].

A further method to increase the speed of analysis is to reduce the number of Raman spectra taken for measurement by highlighting targeted areas for intensive raster scanning, in selective scanning Raman microscopy [155]. This can be performed by predictive algorithm, where spectra are taken at two random points on the sample, and the difference between the spectra informs an algorithm to predict where to take the next measurement [155]. Another method is to use another, less specific, but highly sensitive optical technique such as auto-fluorescence to rapidly assess the sample and determine where to take Raman spectra [156]. This allows a substantial reduction in the number of Raman spectra taken, with a similar diagnostic yield, and in a shorter time period.

Surface Enhanced Raman Spectroscopy (SERS)

SERS is a method to enhance the inherently weak Raman scattering by using receptor targeted metallic nanoparticles combined with bright Raman reporter molecules as biomarkers. Nanoparticles bind to the tissues of interest in a specific configuration and ratio that gives a unique spectrum. These nanoparticles give an intense signal due to their brightness and specificity, and so a large area can be rapidly scanned to give a reliably sensitive and specific Raman Spectrum [157]. However, the obvious disadvantage lies in the time needed for pre-processing of samples with nanoparticles, and the potential risk of toxicity of metallic nanoparticles that may limit its *in vivo* use [158]. The inherent heterogeneity of tumour biochemistry and receptor expression both within and between patients can affect the accuracy of these techniques [159].

Spatially Offset Raman Spectroscopy (SORS)

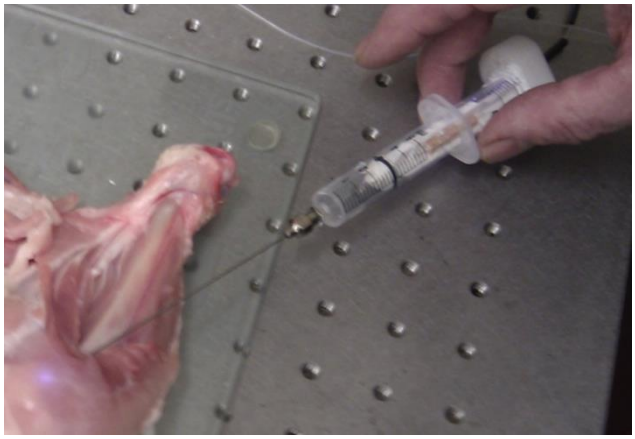
SORS can provide biochemical information at depth below the surface. SORS illuminates at a central point then collects scattered data at a distance from this central point, the light having travelled through varying depths of tissue. In using multi-variate analysis of the resultant spectra, the offset between illumination and

collection is accounted for, and a depth profile of the tissue with tissue information can be gained [160]. An essential component of this technique is that it ameliorates interference by the often- stronger Raman scattering and fluorescence from the tissue surface to be able to analyse the tissue below. The configuration of the laser illumination centrally with an annular arrangement of collection fibres 2 – 3 mm has allowed for analysis 1 – 4 mm below skin [161]. Transmission Raman Spectroscopy (TRS) is an extreme version of SORS, where incident beam and collecting camera are opposite one another, to allow analysis through the sample to gain clinically relevant information at depths of up to 40 mm [162].

High Wave Number (HWVN) Raman Spectroscopy

The majority of biomedical RS uses Near Infra-Red (NIR) light to acquire spectra within the 'Finger Print' (FP) region of 400-1800 cm^{-1} which has been shown to provide extensive detail of the tissue biochemistry. However, when this laser light illuminates the fused-silica fibre required for fibre optic probes it generates an intense background signal [163]. Using the HWVN region of 2400-3800 cm^{-1} can allow Raman spectra to be collected without this background interference [163]. This opens the way to numerous clinical applications and can be used within commercially available devices such as a hypodermic needle [164] core needle biopsy system [165] and endoscopes [166](see Figure 1-17). This advantage must be weighed against the disadvantage of using HWVN region spectra, which is less specific and may have limited diagnostic capabilities compared to data from the FP region [167].

A



B

Figure 1-17 Examples of fibre optic probes capable of Raman Spectroscopy measurement in a clinical setting for cancer diagnosis, or application to in vivo surgical guidance to provide IMA, images Authors own.

A. A fibre optic probe is incorporated into a standard 5ml syringe with a 15cm long, 20 gauge needle and so is capable of subcutaneous measurements (in this example in a Turkey leg) as detailed in reference [168]. B. A miniature confocal Raman packaged probe with a GRIN lens objective for endoscopic use. A detailed review on fibre optics for clinical use of Raman Spectroscopy is found in reference[169]

Coherent anti – Stokes Raman scattering (CARS) and stimulated Raman spectroscopy (SRS)

CARS and SRS are non- linear processes, where the observed effect is not linear to the incident laser power, as it is a result of multiple overlapping photons. It can therefore generate a signal intensity greater than coherent Raman. By probing specific, narrow spectra with high intensity, specific molecular information can be

gained rapidly [170]. The advantages of CARS relate to its sensitivity to CH molecules making it particularly effective at studying lipid and fat distributions, which can been performed at video rate of 100 ns per pixel [171]. However it is a near surface technique and interpretation is made difficult by the presence of a non-resonant background, causing spectral distortion and artefact [170, 172].

A SRS signal is generated when the molecular vibration is equal to the difference in frequency between the pump and Stokes photon. The change in intensity of these beams as a result is measured. Its use was previously limited by slow acquisition times and its reliance on back-scattering meant it was inherently weak. However, new techniques have overcome these issues and allowed rapid acquisition of spectra to give detailed biochemical information [172]. Further, SRS is non-resonant background free, can be performed with ambient light present and its ability to provide high-speed imaging has been used in clinical tissue diagnosis [173].

Spectroscopic Data Analysis

Regardless of the method of Raman spectroscopy utilised to gain spectra, the raw spectra require careful analysis to elucidate clinically relevant information. Differences between tissue spectra, although highly specific and holding detailed chemical information, can be subtle. In the setting of IMA, data analysis is focused on providing a binary outcome of whether the margin is adequate or not. In general, there will be a 'training set', which are spectra assigned to known tissue correlations i.e. a spectrum taken from tissue which has a histopathological diagnosis (as demonstrated in Figure 1-18). These are used to create a model, which is then tested on a 'validation set', where the diagnostic accuracy can be assessed. This training – validation can be run multiple times with a Leave One Out Cross Validation (LOOCV) protocol to provide a measure of independence in the performance measures [107].

The types of mathematical models used to create spectral classifications are numerous, and expanding. A simple technique is Direct Peak Analysis – where individual spectral features are analysed e.g. the areas under an individual Raman band can be compared [174]. Principal Component Analysis (PCA) and Cluster analysis are unsupervised techniques, that do not require assignation of

spectral peaks, but identify where in the spectrum the greatest variance between data lies and classifies data according to these groups. Linear Discriminant Analysis (LDA) is a supervised technique that is commonly used to distinguish differences in the classes identified by PCA to increase accuracy. Increasingly complex models have been developed such as Support Vector Machines (SVM) and Random Forest Classifiers [175] and can improve diagnostic accuracy[176], however they can be more difficult to apply and interpret. The machine learning classifier used to analyse spectral data is an important part of system development. The balance between simplicity and speed of simpler techniques versus the complexity and improved accuracy of more recent models needs to be considered when developing a Raman system suitable for performing intraoperative margin analysis in the clinical environment.

1.3.5 The use of RS to determine the surgical margin

1.3.5.1 Breast Cancer

Breast cancer holds great potential for intra-operative use of RS, with a number of groups reporting a variety of advanced techniques to assess the margins of lumpectomy specimens.

Background

For the majority of patients with early primary breast cancer, Breast Conserving Surgery (BCS) (with adjuvant radiotherapy) offers an alternative treatment to traditional Mastectomy, with equivocal survival rates and improved patient satisfaction (Fisher, 1985). In Breast conserving surgery, the cancer is removed from the breast, termed a lumpectomy. The method of assessing margins are currently mainly by surgeon palpation, if it is a palpable lump, or radiographically with an intraoperative X - ray. Neither are a reliable way of assessing the specimen for involved or close margins [177], and as a result, the rate of close or involved margins is high [91].

The definition of a positive margin is most commonly described as ‘ink on the tumour’ [56, 178], the definition of a ‘close’ margin is much more debated, however in the UK it is defined as < 2mm, however, the resultant practice as to whether a re-excision of margins is recommended, or indeed undertaken, is much more variable [91]. A positive margin is associated with a 2 fold increase risk of local

recurrence, despite adjuvant therapies [92]. Therefore, a positive margin, and more variably, a close margin usually necessitates a re-excision of margins. This is a further operation where the operation site is re-opened and the tumour bed examined, the surgeon then takes a further 1 cm rim of breast tissue at the site that was reported as involved. The number of cases requiring re-excision is high, with a UK average of 17% and some units reporting up to 38-41% [91, 134]. The steps involved in this re-excision are numerous and complex and the sheer volume is a significant financial burden.

RS ability to differentiate normal from abnormal breast tissue

There is a large body of evidence confirming the ability of RS to differentiate between normal, benign and malignant changes. A meta-analysis included 9 studies and concluded that using RS *in vitro* in breast tissue to diagnose breast cancer gave a pooled sensitivity of 0.92 and specificity of 0.97 [179]. However, there was a marked heterogeneity between study techniques, and so studies need to be considered individually.

In 2005, Haka examined lumpectomy and mastectomy tissue from 58 patients that had been snap frozen then thawed. They used Raman micro-imaging to gain a sampled volume of 1mm^3, and determined a total of 130 spectra, which they used in a leave-one-out cross-validation analysis. With this model they got a 94% sensitivity, 96% specificity and an overall accuracy of 86% for detecting infiltrating carcinoma. Fibroadenomas appeared to count for this diagnostic uncertainty; in 2 instances the Raman diagnosis was fibroadenoma, and the histopathological diagnosis was infiltrating carcinoma. The reason may have been the sole differentiation between the two pathologies in the diagnostic algorithm was fat content [180].

In a paper to compare the spectroscopic techniques of fluorescence, diffuse reflectance, combined fluorescence and diffuse reflectance and RS, in the diagnosis of Breast cancer, Majumder *et al.* found RS to be the most effective [181]. They used 74 frozen – thawed specimens to measure 293 spectra with point RS. They found distinct peaks associated with connective tissue proteins and fatty acids discriminated well between normal tissue, fibroadenoma, invasive ductal carcinoma and DCIS being able to classify 99% of spectra correctly. This was

done in laboratory conditions, and the area sampled was not reported, but can be presumed to be small.

Barman *et al.* developed a single step Support Vector Model algorithm using point RS to diagnose breast lesions from 33 patients undergoing Vacuum assisted biopsy, specifically in those with microcalcifications [182]. They achieved an overall accuracy of 82% of diagnosis with excellent Positive Predictive Value (PPV) (the probability of a positive result being a true positive) of 100% and Negative Predictive Value (NPV) (the probability of a negative result being a true negative) of 95% for breast cancer. However, this was performed in physically small biopsy samples, with a selective population of tissue all with calcifications present, which are known to produce relatively intense Raman spectra [183], and thus may be unrepresentative of all breast cancers.

Han *et al.* used a confocal Raman system to look at freezing microtome sections of breast tissue [184]. They defined the peak assignments of Raman spectra associated with breast tissue and found that the relative intensity of the C = O peaks increased with increasing grade of malignancy. They took 475 spectra from 39 patients and identified that there was little inter-subject variation in the spectra. They used a Support Vector Model for their classification model and achieved an overall accuracy of 74%.

An issue with these techniques is that they are surface techniques, so if there was cancerous tissue more than a few hundred microns from the surface this would not be detected, so ‘close’ margins (which usually require excision) of up to 2 mm would go undetected. Spatially offset Raman spectroscopy (SORS) can resolve this and Keller *et al.* developed a SORS probe that obtained spectra at a depth of 2 mm [185]. In 35 samples of frozen- thawed tissue they attained sensitivity and specificity of >94%. They assessed the margins compared to histopathology as simply ‘positive’ or ‘negative’ rather than gaining specific tissue diagnosis. The sampling size of the probe, nor the method of location of tissue sampling is described, but the authors recognize that assessing the entire specimen in a clinically relevant time is a limitation of the technique.

These studies were all done on micro or point spectroscopy, sampling very small areas of breast tissue to obtain Raman spectra. For application to IMA, it is necessary for a large area to be analysed, rapidly and so complimentary techniques have been applied to enable this. Kong *et al.* used auto-fluorescence imaging to inform selective -sampling Raman microscopy to provide an accurate diagnosis within a clinically relevant time frame [156]. Tissue samples cut from blocks that were frozen-thawed were used and sensitivity and specificity of >90% were achieved. An example of these microspectroscopy mapping results and assignation of spectra to tissue is shown in Figure 1-18. This study only differentiated between ductal carcinoma and normal tissue, and in other studies, it is DCIS and fibroadenomas that negatively affect the overall accuracy of the analysis. Once again, the breast tissue samples were small (5 X 5 mm^2), and so the conclusion of analysis within a clinically relevant time frame was extrapolated. How the technique would be applied to a whole specimen without cutting the sample is also unexplored.

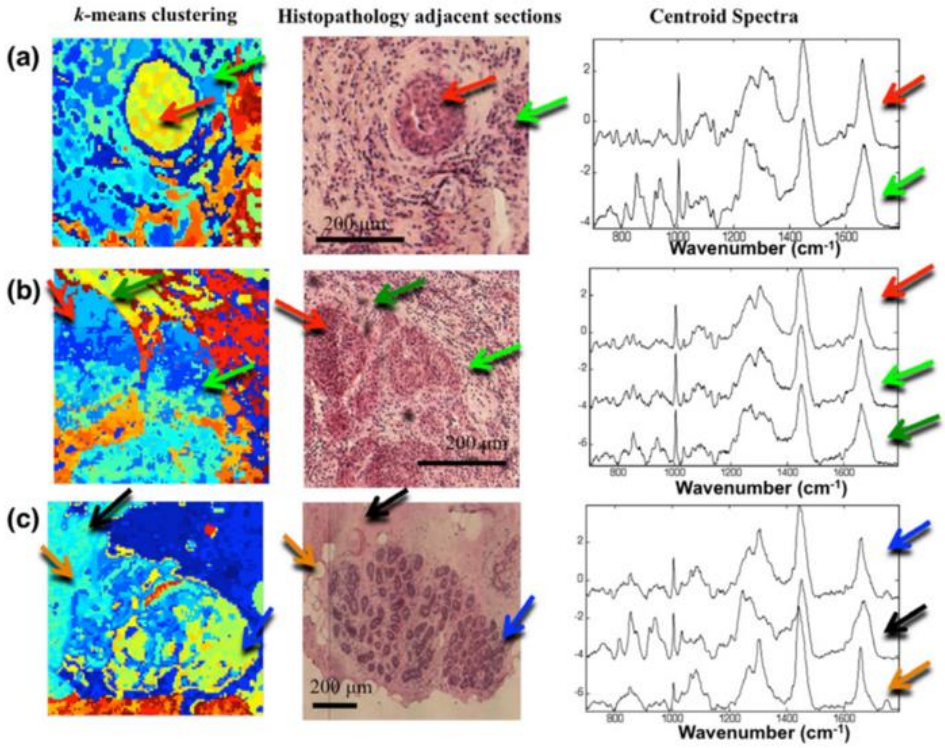


Figure 1-18 Example of assigning Raman spectra to tissue structures and ductal carcinoma of breast tissue to inform the diagnostic algorithm.

(a) + (b) invasive ductal carcinoma (IDC), (c) normal breast tissue. Red arrows show focus of IDC, green arrow tumour surrounding inflammatory stroma, blue arrows lobules and ducts, black arrows, stroma and orange, fat (Reproduced from reference [156] K. Kong, F. Zaabar, E. Rakha, I. Ellis, A. Koloydenko and I. Towards intra-operative diagnosis of tumours during breast conserving surgery by selective-sampling Raman micro-spectroscopy. https://doi.org/10.1088/0031-9155/59/20/6141, under a Creative Commons Attribution 3.0 licence).

Intra-operative use of Raman spectroscopy

Based on these promising initial results, the same groups went on to use RS in the clinical setting with the potential to give an intra-operative diagnosis.

Haka *et al.* used their previously validated technique in freshly excised tissue from 28 patients measuring 220 spectra [186]. Tissue with Ductal carcinoma *in situ* was excluded, as the validating set had not encountered this pathology which is an important exclusion for intra-operative use, as DCIS is associated with a higher rate of re-excision of margins [187]. It also excluded patients having undergone neo-adjuvant chemotherapy, which is increasingly common, and those patients with calcifications. Once again, fibroadenoma proved a diagnostic challenge, and the positive predictive value of 36% can be considered poor, with an overall accuracy of 86%. Although the authors felt that the high NPV of 99% was the main clinically relevant outcome, with such a low PPV in clinical use this would lead to a high volume of breast tissue unnecessarily being excised. The tissue area sampled is not mentioned, meaning the relevance of the technique to assessing an entire sample for margin adequacy is difficult to assess.

Despite these limitations, this was performed adjacent to the operating room, in a light box, and analysis was performed in 30 minutes, recreating conditions necessary for intra-operative use of the technique.

Wang *et al.* used SERS with nanoparticles to assess 57 freshly excised lumpectomy specimens and processed samples within 15 minutes [188]. Each specimen was topically stained with Raman active nanoparticles that were functionalised with antibodies to target HER2, mER, EGFR and CD44 and then raster-scanned to acquire spectra for the entire resected surface on the exposed glandular tissue. It was possible to differentiate between normal, benign changes and invasive carcinoma tissue, and the overall sensitivity for breast carcinoma detection was 89% with 92% specificity, with the accuracy for the specific biomarkers slightly less than this. This technique is not affected by haemoglobin, surgical dyes or diathermy increasing the clinical relevance, however, it is limited by the sensitivity and specificity of not only the functionalised nanoparticles ability to bind to the molecules of interest, but also the accuracy of the Raman signal of the nanoparticles. Another limitation is that, as a surface technique, ‘close’ margins are not detected.

Using a method of selective scanning, Shipp *et al.* performed analysis on freshly excised lumpectomy samples from 51 patients, and analysed one resected surface of each specimen which was identified as being most likely to be positive by a member of the team [189]. They used multimodal spectral histopathology (MSH), obtaining autofluorescence images, which was highly sensitive but non-specific, to inform targeted Raman measurement points in identified 'segments' to reach a final diagnosis. The diagnostic algorithm was validated using a training set based on mastectomy samples which included tissue containing fibroadenoma, fibrocystic change, DCIS and invasive carcinoma. MSH in the lumpectomy samples was highly sensitive, identifying all the margins that contained residual cancer even as small as 1 X 1 mm^2, and so was 100% sensitive, with around 80% specificity. They analysed a single resected surface of up to 4 X 6.5 cm^2 in 12-24 minutes, which was achievable as MSH reduced the number of Raman spectra required by 100 – 200 fold. Although this study shows significant improvement in the speed of analysis and the diagnostic accuracy, there are some limitations in the way the lumpectomy specimens were subsequently examined, and that only one resected surface could be examined in a clinically relevant time.

1.3.5.2 Skin Cancer

Background

Skin cancer is the most common cancer diagnosed and its diagnosis and treatment represents a huge burden on the health economy [190]. Basal Cell Carcinoma has the highest incidence and is predicted to rise. It is a slow growing tumour, that rarely metastasises, but local invasion leads to local tissue destruction and deformity. Surgery remains a treatment of BCC, the tumour can be excised by standard surgical excision where the lesion is excised with the aim of obtaining margins of 3 – 5 mm. A positive, or close margin has been reported in up to 7% of cases and is associated with a high recurrence rate of up to 27% [191] and requires further treatment or re-excision [192]. Mohs' micrographic surgery is a current technique used for IMA but it is time consuming, expensive and only recommended in high risk cases [192, 193]. Melanoma is less prevalent than non-melanotic cancer but incidence is rapidly rising, it has metastatic potential and late presentation is associated with a very poor prognosis [194]. The management

of melanoma is based on surgical excision, with margins of 0.5 – 2 cm required based on the stage of disease, and re-excision performed if there are involved margins[195]. A margin narrower than this is an inadequate resection which increases the risk of recurrence associated with a poor prognosis, and may possibly be associated with worse survival [196]. For both melanoma and nonmelanoma skin cancer, there is a need for IMA to improve patient outcome.

RS ability to differentiate normal from abnormal skin tissue

Confocal RS was first used to differentiate BCC from normal skin in 2002, where Raman maps from a small sample of 15 specimens were taken and compared to surrounding normal tissue which yielded sensitivity and specificity of over 90% in a logistic regression model [197]. Acknowledging the practical limitation of using confocal RS, the group used a handheld fibre optic probe and gained spectra using HWVN RS in the region of 2500 – 3800 cm^{-1} to avoid background signal from silica [198]. A number of readings from each of 19 biopsies taken from the centre of confirmed BCC's were analysed, which demonstrated large and consistent differences between the spectra from BCC and normal tissue, particularly that collagen contains discriminative information at this wavelength, with a 100% diagnosis of BCC. However, gaining spectra at this wavenumber region took longer due to suboptimal signal-to-noise ratio, which may limit its clinical use, and there was a small study size.

These studies looked solely at BCC, whereas non-melanotic skin cancer also includes squamous cell carcinoma. Lieber *et al.* analysed 21 suspected non melanotic skin cancers and took measurements with confocal RS from within the tumour and compared to normal skin adjacent (1 cm distant) from the tumour edge at a depth of 20 μm [199]. They achieved good sensitivity of 100% for determining the cancerous lesion and specificity of 91%, with squamous cell carcinoma lesions accounting for the diagnostic uncertainty. The sample population of 19 patients was small and although marked differences in Raman spectra were observed for each pathology this was after subtracting the matched normal reading. There is marked inter-subject variability in the Raman spectra of normal skin, and without a matched normal reading the diagnostic accuracy is likely to have been affected. The same group went on to perform measurements at varying depths on fresh-frozen samples from 39 patients with no skin

pathology, BCC, SCC or melanoma and achieved 100% diagnostic accuracy at the surface which decreased towards depths of 100 μm [200].

Schleusener *et al.* recruited 104 patients scheduled for excision of a suspicious lesion and used a fibre coupled probe *in vivo* with direct skin contact to sample 3 measurements on each lesion and the mean used to inform the spectra [201]. The heterogeneity of the lesions led to poor accuracy in determining non-melanotic skin cancer from normal skin cancer. The major differences in skin variability between body site also contributed to the results that achieved an accuracy of 78% in discriminating BCC and SCC from normal skin. For melanotic lesions the lesion inhomogeneity was insignificant, and they achieved a balanced accuracy of 91% of differentiating melanoma from normal pigmented nevi.

In a large study, Lui *et al.* investigated all suspicious skin lesions *in vivo*, both potential non-melanoma and melanomas, in 848 patients and acquired 1022 spectra [202]. Spectra from the *in vivo* lesion were taken, and compared with spectra from normal appearing skin 5 cm from the tumour edge. The aim was to detect which lesions required invasive biopsy to histologically confirm malignancy which was achieved with 90% sensitivity and 64% specificity. The strength of this study was its clinical relevance – measurements were gained within 1 second, with a handheld probe, and was concerned with the relevant task of differentiating cancerous lesions from benign lesions, rather than from normal skin like other studies. However, the results were ultimately disappointing, with poor specificity. This may have been due to the heterogenous group of benign lesions to compare against (which didn't necessarily have a confirmed histopathological diagnosis), and a relatively small number of malignant melanomas (n=44) which may have underpowered the diagnostic algorithm.

The same group used a similar approach with a probe measuring a diameter of 3.5mm at a depth of 1mm to validate the previous diagnostic algorithm on a new cohort of 127 cases, where they achieved similar sensitivity and specificity in cancer versus non cancer diagnosis to the previous study[203]. It is noted that, setting the sensitivity level to 95%, only 8 of 9 melanoma cases were correctly classified as a cancer, and the specificity was generally poor at 30 – 46%

depending on the sensitivity level. The results are perhaps unsurprising as the diagnostic algorithm had previously had poor accuracy at melanoma diagnosis.

In a meta -analysis to investigate the accuracy of RS for differentiating cancer from normal tissue, 12 studies using different methods of RS were included, then analysed according to *ex vivo* and *in vivo* studies and types of skin cancer. 10 of the studies investigated discrimination of BCC, and *in vivo,* the pooled sensitivity for discriminating BCC from normal tissue samples was 69% and specificity 85%, compared to *ex vivo* sensitivity of 99% and specificity 96% [204]. This suggests the use of RS to examine *ex vivo*, resected cancer samples, for margin assessment may be a highly accurate technique.

There are a number of studies investigating mixed methods of combining RS with other techniques to improve performance. Combining autofluorescence with RS, with six spectroscopic criteria, 79 *ex vivo* patient samples were analysed and cancerous tissue was classified with 97% accuracy [205]. Another group used CARS, second harmonic generation and two- photon excited fluorescence imaging to analyse 140 *ex vivo* skin samples in a multimodal approach that allowed large -area scans and the identification of dermal layering, which may assist in diagnosis of cancerous lesions [206].

Intra -operative use of Raman spectroscopy

Kong *et al.* used a method of measuring tissue autofluorescence to determine the sampling points for RS [207], a technique alluded to in the previous section [208]. This MSH technique was used to analyse samples from 20 patients treated with Mohs' microscopic surgery for BCC, half were BCC positive. Analysing tissue samples of up to 1 X 1 cm^2 the sensitivity and specificity was 95% and 94% respectively for the detection of BCC within a time of under 60 minutes. The same group has now reported a fully-automated prototype instrument based on this technique that allows assessment of skin surgical resections of up to a 2 X 2 cm^2 area which detects residual tumour at the surface of the resected sample [209]. This prototype will be used to validate their previous work with a larger population of patients.

1.3.5.3 *Brain cancer*

Background

Gliomas are tumours of the neuroglia graded by histopathological features that account for the majority of malignant brain tumours in adults. They have varying prognoses, but the commonest, glioblastoma multiforme accounts for 55% of all gliomas and has a 5 year prognosis of 5% with almost inevitable recurrence after treatment [210]. Surgical resection is generally recommended as part of initial management for both histological diagnosis and to remove as much of the tumour as possible, if it is safe to do so [211]. A major challenge of resectional surgery is achieving adequate margins, particularly as excessive tissue excision can lead to major neurological morbidity for the patient. With current imaging techniques of neurosurgical microscopes or intraoperative MRI, even in cases of an apparently complete resection with 'clear' surgical margins, the vast majority of recurrences occur at the site of resection , suggesting current techniques of assessing intraoperative resection are inadequate [212].

Improving the intra-operative assessment of surgical resection margins could improve adequacy of tumour resection, and thus recurrence rates for glioblastoma.

RS ability to differentiate normal from abnormal brain tissue

The majority of initial diagnostic work has been performed in mouse models of brain tumours. RS has been used to analyse tissue from mouse models of glioblastoma *ex vivo* and is able differentiate between normal tissue (white and grey matter) and malignant tissue with 100% accuracy [213] and has been used to examine the tumour margins in mice *in vivo,* where RS identified tumour undetected by bright field microscopy [214]. Uckerman *et al.* used CARS to probe the C-H molecular vibration, thus imaging the lipid content of samples [215]. A mouse model of glioblastoma was analysed *ex vivo* then the same technique used in human glioblastoma tissue to confirm the findings. They found malignant tissue was identified by a reduction in lower CARS signal intensity which was related to a lower content of total lipids in tumour tissue than normal tissue. This was at a cellular level and so tumour borders could be discerned precisely, the technique could gain images at 20 Hz, representing clinically relevant time for intra-operative use.

Two studies using induced glioma formation in mice models have reported the use of systemically injected gold nanoparticles preferentially up- taken by tumour to inform SERS guided tumour resection [216, 217]. The nanoparticles are hypothesised to cross the Blood Brain Barrier via low-density lipoprotein-receptor -related protein 1, an active transport endothelial receptor that carries exogenous substrates across the BBB [218]. The activation of the nanoparticles is then reliant on the acidic tumour environment, which results in a specific signal at the tumour site [216]. A handheld Raman probe was used in both cases to demonstrate the delineated tumour margin and guide tumour excision. Although they show promise in mice models at assisting in obtaining clear excision margins, the translation of using a systemic agent in humans for diagnostic purposes only is likely to be complex and remains un-investigated.

In human tissue, Raman microspectroscopy has been used to differentiate normal brain tissue, necrosis and brain metastasis and achieved accuracy of >95% [219]. Kalkanis *et al.* used *ex vivo* human tissue from 17 donors to create histological slides from frozen samples. Within homogenous areas of normal, necrotic and glioblastoma areas a diagnostic accuracy of over 97% was achieved in the validation group. However non-diagnostic areas, heterogenous areas and those with freeze artefact were excluded which limited sample size, and limits validity of application *in vivo* where heterogenous areas are likely [220].

Another clinically relevant study by Ji *et al.* used Stimulated Raman Scattering (SRS) to examine tissue from 19 patients with CNS malignancy [173]. They produced two- colour images based on the Raman intensity ratios which displayed whether the structure was lipid or protein-rich. Using biopsy samples, they asked pathologists to compare these SRS images with standard H+E pathology images and achieved excellent diagnostic concordance. This was carried out in standard lighting conditions. However, they acknowledge the areas sampled were much smaller than a true tumour bed, and SRS can only sample at a depth of 100 μm. Furthermore, this approach is only likely to work when normal tissues are mostly lipid rich and the cancer or disease leads to a change in protein rich tissues. Another group using SRS to analyse histology slides achieved similar results in differentiating between normal brain tissue and tissue containing a lesion. They used this to develop a machine learning process that

was able to predict an automated diagnosis of tumour subtype with 90% accuracy [221].

Intra -operative use of Raman spectroscopy

A recent study by Bury *et al.* analysed 29 fresh brain tissue samples that had been excised during surgery within a clinically relevant time[222]. Using SERS, the samples were processed with gold nanoparticles and Raman spectra obtained, the tissue then underwent routine histopathological analysis. There were a number of diagnoses within the small sample size meaning it is likely to be statistically underpowered. Despite this they gained sensitivity and specificity above 75% in diagnosis of normal, glial and metastatic brain tumours, with meningiomas proving a diagnostic challenge with poorer accuracy. Results were comparable to currently used methods of IMA and superior diagnostic accuracy is needed for clinical adoption. However, this could be overcome by increasing sample size and measurements were taken in real time in a laboratory linked to the operating theatre via air-tube, which is an innovative solution to the often encountered problem of space, and demonstrates successful clinical application.

In 2015 Jermyn *et al.* reported the use of a handheld spectroscopy device that used a Raman fibre optic to perform sub-millimetre single point measurements of 0.2 mm^2 *in vivo* in humans [223]. 161 MRI guided intra-operative measurements were taken from 17 patients with gliomas, and a biopsy taken at the corresponding site for correlation of Raman spectra with H+E pathological diagnosis. They found specimens with cancer cells had a difference in the lipid bands, a higher nucleic acid content, and an increase in the band associated with the breathing mode of phenylalanine in proteins. Tissue with cancer cells present were distinguished from normal brain tissue with an accuracy of 92%, which was significantly better than the operating surgeons' visual analysis with a bright field microscope. This was performed intra-operatively with a small, hand held probe and measurements took less than 1 second. The limitations related to the restricted field of view offered by the small area sampled by the probe, and the false negatives in the Raman analysis were due to the system needing > 15% cancer cell burden to be accurate. The same group then went on to integrate intrinsic fluorescence spectroscopy, diffuse reflectance spectroscopy and RS into one system to analyse biopsies taken from 15 patients with brain tumours of any

type in a similar study design [224]. Using this multimodal approach they achieved sensitivity of 100% and specificity of 93% in differentiating between normal brain tissue and tissue with cancer cells.

This group have recently developed a probe incorporated within a commercially available biopsy system to allow Raman measurements to be taken without disrupting surgical workflow [165]. It used HWVN RS to collect data mainly from lipids and proteins. It was successful at detecting normal brain tissue and dense cancer tissue but could not differentiate between normal brain tissue and tissue infiltrated with cancer- which is likely to represent the specimens with inadequate tumour resection margins.

1.3.5.4 Head and Neck Cancer

Background

Head and neck squamous cell carcinoma (HNSCC) represent the main oncological burden of head and neck oncology. Resection remains the mainstay of treatment for the majority of HNSCC locations [225]. Complete resection of the tumour is the goal of surgical treatment, as a positive margin doubles the risk of local recurrence compared to those with a negative margin [226]. Despite this goal, a significant proportion (30-65%) of HNSCC resections have positive resection margins [227]. A pathologically involved or close margin affects further management which is often the use of adjuvant therapy such as chemotherapy and/or radiotherapy. Re-resection can be considered, but only if anatomical location allows and after associated morbidity is considered [124, 228]. A common definition of a close margin is <5mm for HNSCC [124, 229]. The Intra-operative technique for margin analysis has been frozen section which has been specimen or tumour – bed based, with variability in the way it is utilised, with no standard method adopted [230]. However, there has been no convincing evidence that this reduces the positive margin rate or improves outcome [128, 231].

RS ability to differentiate normal from abnormal head and neck tissue

The first report of RS to differentiate between normal and malignant larynx tissue was from Stone *et al.* [232]. Raman microspectroscopy was used on biopsies from 19 patients to differentiate homogenous samples of normal tissue from dysplastic and squamous cell carcinoma tissue. Diagnostic peak height ratios were used

rather than absolute spectral peaks to inform a diagnostic prediction model that demonstrated sensitivity and specificity of >90% for the diagnosis of squamous cell carcinoma.

Using frozen – thawed biopsy samples of vocal cord lesions, Lau *et al.* analysed areas of 3.5 mm within 5 seconds. There was heterogeneity of tissue type within measured samples which may have accounted for the reduced diagnostic sensitivity of 69% of detecting carcinoma [233].

Lin *et al.* developed a miniaturized RS fibre optic probe that was inserted down a working channel of nasendoscopy instrument to take measurements of suspicious laryngeal lesions in 39 patients [234]. The probe was put in contact with the lesion for < 1 sec prior to taking biopsies for histopathological analysis. They used the HWVN range (2800–3020 cm^{-1}) to obtain 94 spectra and identified spectral peaks that successfully differentiated normal and malignant tissue. In a similar experimental design in 60 patients, the same group used a combination of FP and HWVN RS to acquire over 2000 spectra and compared this to histopathological biopsy. They gained spectra rapidly in < 1 second, and the combined spectra yielded an overall diagnostic accuracy of 91.1% [235]. The same group then acquired spectra from 90 patients with nasopharyngeal cancer and used PCA and LDA with a leave-one-out cross validation method to achieve a diagnostic accuracy of 93.1% [236].

In the detection of oral carcinoma, Cals *et al.* took histological sections from 11 samples of oral SCC with surrounding normal tissue, and histopathological evaluation then selected the regions for RS measurements [237]. Raman mapping with an automated confocal Raman microscope took point measurements at 5 μm steps to determine spectral differences between oral SCC and squamous epithelium, connective tissue, gland, muscle, adipose tissue and nerve. They achieved excellent distinction between SCC and healthy tissues with >97% accuracy. They went on to develop a two-step classification model using a similar experimental method for validation in 19 samples and achieved diagnostic accuracy of 91% to differentiate tumour vs non-tumourous tissue [238].

In vivo detection of oral lesions was performed by Guze *et al.* with a handheld probe in real time [239]. The probe, which had disposable plastic sleeves, was used to diagnose lesions within 5 minutes and the procedure was tolerated well by the 18 patients who had a previously known histological diagnosis of the oral lesion. They were able to differentiate between pre-malignant and malignant tissue versus normal or benign lesions with 100% sensitivity and 77% specificity.

The difference in water content between normal tissue and SCC has been used by one group as a marker to identify cancerous tissues in the head and neck. Using a confocal Raman microscope with HWVN RS at 2500 - 4000 cm^{-1} they used freshly excised oral SCC samples from 14 patients to take up to 30 spectra from each sample within 30 minutes and subsequently compared them to histopathological evaluation. They found the intensity of the OH-stretching vibration increased in SCC more than normal tissue, along with the water concentration being significantly higher in the SCC containing tissue. They concluded that water concentration could be determined with HWVN RS and was a useful diagnostic marker of SCC tumour [174]. The group then used freshly resected oral SCC specimen sections containing both tumour and normal tissue to analyse how water concentration changes with distance from the tumour. Using a confocal Raman microscope at the same wavelength they then obtained over 3000 Raman spectra to determine that mean water concentration within the tumour was 76% and decreased further away from the tumour down to 54% when > 4mm from the tumour in healthy tissue[228]. A similar design was used in 26 mandibulectomy specimens and it was also found that water concentration is high within tumour (mean of 77%) and decreases with distance from tumour to a mean of 44% in healthy tissue. These spectra were then used to develop a classification model for diagnosing SCC in bone in a training set and in the validation set achieved a sensitivity of 95% and specificity of 87% in tumour detection [240]. They showed good diagnostic accuracy within clinically relevant times of less than 30 minutes. However, there are limitations to clinical applicability; the specimens had to be handled in a particular way to avoid desiccation, and these studies used cut specimens which may have different water properties to an uncut specimen. A flat surface was also necessary to achieve adequate contact with the Raman probe, which may not be achievable with a clinical specimen.

The data demonstrates a large inter-subject variation in water content in healthy tissue of 17% in bone and 24% in mucosal tissue which suggests pathological tissue measurements would always need to be compared with concurrent 'normal' tissue readings [240]. The calculation of water content for these measurements were based on a protein model and ignored the contribution of lipids to the measured spectral band. Although this is acknowledged and in oral mucosal tissue any high lipid signals were always associated with healthy tissue, this may not be the case with other tissues and so the ability to apply this methodology to other pathologies may be limited [240].

1.3.5.5 Other pathologies

There has been investigation into RS in the diagnosis of other solid tumours such as ovary [241], lung [242, 243] and thyroid [244], but with little further exploration to the application of this technology to improving the adequacy of surgical excision margins. In some solid tumour pathologies, the use of RS *in vivo* for the detection of cancer for identification of residual tumour and ensuring adequacy of resection is another method of improving surgical oncological outcomes.

In prostate disease, histological studies have differentiated between benign prostatic hyperplasia, prostate cancer and normal prostate tissue with a sensitivity of 94% and sepecificity of 100% [245], which was seen to be due to increases in DNA and collagen changes in malignancy [246]. In freshly excised tissue both the Fingerprint and HWVN region has been used with a hand held probe to get diagnostic accuracy of over 85%[107, 247]. The same group have succesfully integrated a RS probe into the arm of a robotic operating system and demonstrated the possibility of *in vivo* surgical guidance to ensure there is no residual disease [248].

There has also been substantial research into the detection of early malignant change in the cervix, aided by the well- defined nature of the disease and ability to gain measurements without excision of tissue. Multiple studies have demonstrated the ability of RS to differentiate between colposcopically normal and abnormal areas of cervical tissue to a clinically relevant degree of accuracy [249-251]. This may be useful in improving early, accurate diagnosis to guide targeted treatment and ensure complete resection of any cervical pre-cancers.

The ability of RS probes to be incorporated into fibre optics has significant benefit to the potential for use as a surgical adjunct. In bladder cancer fibre optic RS probes have been shown to be able to differentiate normal bladder and bladder cancer with an accuracy of 84% in pathological samples [252]. Another group developed a fibre optic probe used *in vivo* to gain measurements at sites within the bladder of 32 patients with suspected bladder cancer that were subsequently biopsied and then matched with the definitive histopathological diagnosis. These measurements took place alongside fluorescence cystoscopy, a technique already in use to improve bladder cancer detection compared to simple white light cystoscopy. This clinically relevant technique obtained a sensitivity of 85% and specificity of 79% [253].

The introduction of RS enabled fibre optic probes down working ports in endoscopic instruments holds promise for early diagnosis of oesophageal, gastric and colorectal pathology, with *ex vivo* and *in vivo* studies demonstrating consistently good diagnostic accuracy and clinical relevance [254-257]. The utility of this in achieving adequate surgical margins has not been investigated but given the increased risk of recurrence associated with involved circumferential resection margins in GI cancers [258-260], further work should be considered.

1.3.6 Translating Raman to the clinical environment for IMA

The ideal method for providing IMA would be highly sensitive, not alter the specimen (to allow subsequent histopathological analysis), recordable to allow accountability, give a definitive answer which is easily understood without the need for specialist training, and processes the sample without delaying surgery. The tool to deliver the analysis should be portable to allow use in multiple locations, robust to withstand everyday use, easily sterilised and not interfere with the theatre environment / procedure [261].

In many ways RS meets these criteria. The ability of RS to differentiate between cancer and normal tissue in a non destructive manner has been established and appears reliable in a range of pathologies. The practical advantages of Raman as a tool for IMA are the relatively cheap equipment (£10-30K) is also small and transportable allowing for easy insertion into operating theatres. RS systems

have been developed that have taken measurements within the confines and limitations of the clinical environment, overcoming the obstacles of theatre lighting, handling the specimen and the need for a disposable/re-sterilisable component [224, 262, 263], demonstrating its ability to perform in the operating theatre.

Despite these promising advances, RS is still not used in routine clinical practice, suggesting limitations to the technique for providing IMA. For effective translation and widespread adoption, Raman systems must be cost-effective. The advances in detector technology and lower cost lasers resulting in cheaper Raman systems is addressing some of the historical short falls in the technique, though it is now the detectors and cameras that account for the greatest expense [264], systems able to make use of high-quality but mass produced CMOS cameras used in mobile phones may be expected to reduce costs significantly.

The time taken to analyse samples remains an issue, however innovative techniques such as selective scanning microscopy and SERS has reduced this time and studies presented in this review analysed samples in a clinically relevant time frame of 15 – 60 minutes, which is within acceptable and clinically relevant limits [185, 188, 189, 208, 265]. The fact that sample analysis can take place within the operating theatre obviously saves significant time compared to techniques that require the specimen leaving theatres, such as frozen section analysis.

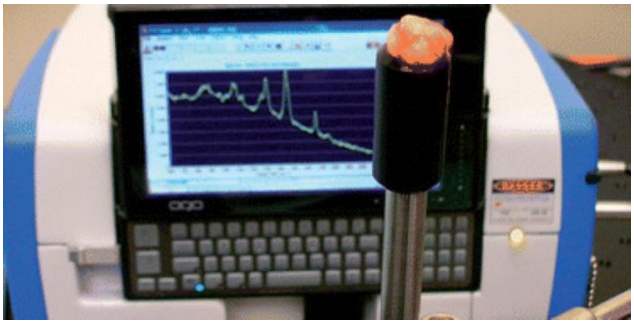
Thus far, studies have required large data sets with complex and potentially lengthy chemometrics to provide accurate diagnostic information. Generally academic teams have been gaining spectra for a training set to construct a diagnostic algorithm. This process requires a significant amount of data processing and handling with large volumes of data and computing power to ‘train’ the diagnostic algorithm, which would be well beyond the capabilities of routine clinical staff. However, these are preliminary studies, where the diagnostic algorithm is being constructed and tested, but once the diagnostic algorithm has been refined, robustly tested and validated, the data produced from a single specimen for analysis would not be overwhelming. The speed of running new measurements through a pre-constructed diagnostic algorithm is in the order of milliseconds.

When planning for the translation of a technology to the clinical environment, the focus must be on the end -user, which for IMA will be surgeons, as they will ultimately determine whether the technique is adopted. When using new technologies there can be difficulties with inter-user variability- in one study assessing IMA in breast specimens using a bio-impedance spectroscopy probe, the results were negatively impacted by the surgeons incorrectly following the probe protocol or incorrectly interpretating the results [266]. Inter-user variability may prove particularly problematic for hand held probe systems, where data can be rejected or inaccurate due to incorrect probe positioning [185, 201]. However, there are a range of other systems such as an automated tissue processing machine that uses cassettes[209], or automated 3D scanner [265] that may reduce this potential for user error.

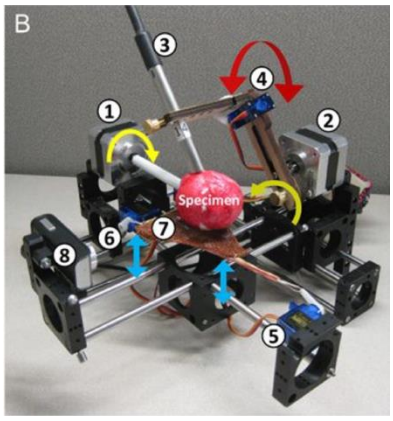
Not only must the measurements be taken by surgeons, but a clear and definitive interpretation of data is required in order to translate to clinical use. It cannot be expected that surgeons should be required to understand and interpret raw Raman spectra to inform the procedure. Examples of systems that provide an indicator of the Raman IMA result to the surgeon is the system capable of emitting a sound to indicate abnormal tissues [267] and the automated 3D scanner 'Marginbot' which has the potential to analyse a specimen and provide a colour coded interpretation of the Raman spectra for the surgeon mapped onto the 3D image of the specimen (see Figure 1-19 B) [265].

So it can be seen that despite the inevitable challenges in translating from bench to bedside (or theatre-side), there are solutions that enable clinical Raman systems to provide easily interpretable IMA of surgically excised tissue to aid intra-operative decision making. Figure 1-19 provides examples of existing Raman probes that have the potential to provide IMA, demonstrating the beginnings of successful translation into the clinical environment. Indeed, there are a number of companies developing commercially suitable Raman systems showing a move away from the laboratory and towards larger scale use by clinicians [264, 268].

A



B



C

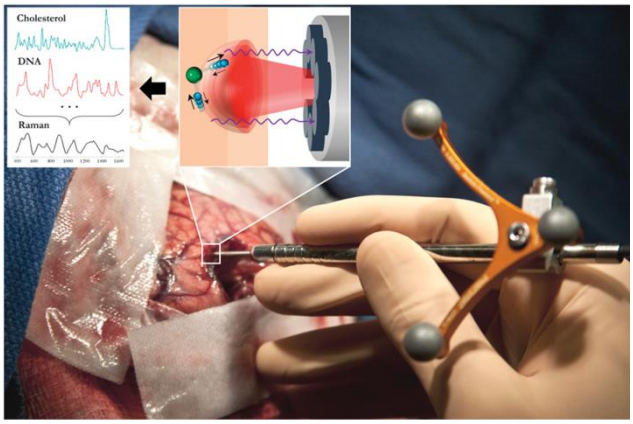


Figure 1-19 Examples of Raman probe systems that could be used for IMA using Raman spectroscopy.

A. Ex vivo RS analysis of a specimen, where the specimen is placed on top of a probe to enable a surface to be analysed. This example uses an axillary lymph node (reproduced from reference [269], with permission from the Royal Society of Chemistry, and the authors) B. Design of an automated 3D margin scanner prototype (Marginbot), where the specimen is placed on a stage and automated movement of the specimen and the Raman probe (depicted by arrows) is required to assess the specimen margins (reproduced from reference [265] G. Thomas, T. Q. Nguyen, I. J. Pence, B. Caldwell, M. E. O'Connor, J. Giltnane, M. E. Sanders, A. Grau, I. Meszoely, M. Hooks, M. C. Kelley and A. Mahadevan-Jansen, Evaluating feasibility of an automated 3-dimensional scanner using Raman spectroscopy for intraoperative breast margin assessment, https://doi.org/10.1038/s41598-017-13237-y, under a creative commons attribution 4.0 International License) C. Handheld probe (Emvision, LLC) for use *in vivo*, in this example to interrogate brain tissue during surgery with the potential to assist in gaining clear margins in the excised specimen. The schematic diagram illustrates the excitation of different molecular species that produces a Raman spectra. From M. Jermyn, K. Mok, J. Mercier, J. Desroches, J. Pichette, K. Saint-Arnaud, L. Bernstein, M.-C. Guiot, K. Petrecca and F. Leblond, *Sci Transl Med*, 2015, **7**, 274ra219-274ra219 [223] . Reprinted with permission from AAAS.

1.3.7 Summary

This review has outlined the importance of the optimal management of surgical margins for oncological excised specimens, current methods of IMA and a review of the existing literature relevant to the use of RS in IMA in a number of solid organ tumour pathologies. It must be noted that the majority of RS studies remain in the realms of laboratory work, or 'processing labs' adjacent to theatres with the work performed by members of academic units. Protocols have not evolved to the point of being able to be used by non-experts, which is crucial to its expansion into the clinical arena. Other disadvantages of RS is the time taken for spectral acquisition, though this is being addressed by multimodal techniques [189], using the HWVN spectra [240] or automation of specimen processing [265]. Ultimately, larger scale clinical studies are required to demonstrate the diagnostic accuracy of the technique, and subsequent improvement in patient outcomes. As part of this, probes suitable for regular clinical use will have to be developed and go through the relevant regulatory processes, and inevitable cost-effectiveness evaluation. The focus on translation of RS to the clinical environment must persist. In an increasingly competitive market of emerging disruptive technologies, future studies must focus not only on improvement of outcomes compared to established techniques of IMA, but also show relevance amongst novel technologies and techniques.

Despite these hurdles, RS has the ability to provide detailed biochemical information of surgical margins with excellent diagnostic accuracy in a range of solid tumour pathologies. Further studies are necessary for the translation of this technology to a clinically relevant environment and demonstrate improved patient outcomes. RS techniques have the potential to provide intra-operative margin analysis of surgically excised solid tumours.

1.4 The role of water content in breast cancer diagnosis

1.4.1 Introduction

The previous section highlighted the work performed by groups in using RS to assess the intraoperative margins of breast cancer specimens. Although the diagnostic ability of RS to differentiate between normal and cancerous breast tissue is good, a major limitation is the time it takes to analyse the large area that would be necessary for IMA. Other groups have investigated using nanoparticles and SERS [270] and multi-modal spectral histopathology combining tissue auto-fluorescence with FP RS [189] as techniques that may reduce this time.

High Wave Number Raman Spectroscopy (HWN RS) is an increasing area of interest in the analysis of biological tissue. The spectral features in this region are from CH-stretching vibrations from protein and lipid at 2800-3040 cm^{-1}. OH stretching at 3100-3500 cm^{-1} which is primarily due to water with some contribution from the NH stretching vibrations between 3100 and 3500 cm^{-1} [103].
The time taken to analyse specimens is reduced using the HWN region due to the reduced volume of data, more intense signals [271], and subsequent simplified data analysis. In addition to this clinical advantage of reduced time for analysis, there is evidence that the diagnostic yield gained from this part of the spectrum is similar to that of the FP region [272, 273]. HWN RS has been used to differentiate between normal and cancerous cervical tissue [272], oral tissue [174], bone [240] and brain tissue [165]. In these tissues, a difference in water content has been a consistent discriminatory feature in the diagnostic analysis, with cancerous tissue having a higher water content than normal tissue.

1.4.2 Theories to explain differences in water content in cancerous tissue

Water content was first noted to be higher in cancerous tissue compared to the surrounding normal tissue in 1971, in rat tissue with sarcoma [274]. Despite this being noted decades ago the biological mechanism as to why water content is higher in cancerous tissue compared to the surrounding normal tissue has not been fully elucidated. One suggestion is that with the increase in cellularity in tumours, there is a reduction of the apparent diffusion coefficient as measured by MRI [275]. With this reduction in apparent diffusion coefficient in cancerous tissue the diffusion of water slows down and so water is trapped [276]. These changes in

the water diffusion may be related to the increased expression of Aquaporins in cancerous tissue, which allows water influx [277].

In measuring the broadband diffuse optical spectroscopy (DOS) of breast tissue, Chung *et al.* [278] calculated that the bound water index decreased in tumours, suggesting an increase in free water in tumours, which could be due to increased levels of hyaluronic acid and necrosis within tumours. However neither hyaluronic acid nor tissue necrosis were actually measured in this study, but it has been shown that an increased rate of cell necrosis is correlated with an increase in water content [279].

Specifically in breast tissue, a likely reason for the difference in water content is because normal breast tissue is predominantly fatty and hydrophobic, and cancerous tissue is predominantly stromal with a large amount of protein, which is associated with a higher water content [280].

1.4.3 Water content in normal breast tissue

The water content in normal breast tissue has been estimated using Diffuse Optical Spectroscopy (DOS), which is a non-invasive optical technique that measures near infrared absorption and scattering. Tromberg *et al.* performed DOS in line scans *in vivo* on 12 patients who were young (aged 30 – 39), and it was demonstrated that the mean water % was 27 (+/- 12 SEM). There were limitations with the sampling volume of the line scanner used and the sampling depth was only 10mm from the skin surface, and therefore these figures for water content may not be representative of the deeper breast tissue water content. Using Near Infra Red spectroscopy in the healthy breasts of 24 females it was found that the water percentage in breast tissue to be mainly between 40 – 60%, however the raw data suggests a wider range was measured from 20-80%, with an acknowledged error range of almost 10% suggests that with such a wide range, and small number of subjects, a firm conclusion cannot be drawn [281]. A study that examined differences in the breast density as measured by MRI measurement of water content found a significant difference between the median water content of mothers (27%) and daughters (44.8%), but again examining the raw data it suggests the range was wide and overlapping, as there were women in both groups with a range of breast water content of between 20 to 80% [280].

This suggests there is significant physiological variation in the water content in normal breast tissue. A number of factors may affect the water content of breast tissue:

- *Age*
    - Independent of menopausal status, as women get older, the breast gets more fatty, there is less variation in lipid measurements, and a corresponding decrease in water content [282, 283]
- *Menopausal status*
    - Premenopausal breasts have a higher water and low fat content compared to post menopausal breasts [282-284].
- *Relationship to menstrual cycle*
    - It was demonstrated with repeated MRI measurements at different times in the menstrual cycle there were cyclical differences in the water content of the breast, however the clinical significance was unclear [285]
    - Using NIR imaging in 7 premenopausal patients there was variation in water content in radiographic density category, with absolute differences of up to 30% of water fraction in normal tissue throughout the menstrual cycle [286].
- *Use of oral contraception*
    - Those patients taking oral contraception had breasts with higher water content and lower lipid content than those not on oral contraception [283]
- *Location in breast*
    - There are differences in water content not only between the two breasts of the same patient, but also within the breast and where the measurements were taken, with areolar water concentration being higher [282].
- *Body Mass Index*
    - As BMI increases, the water content decreases and lipid increases, as the breast is composed more of fatty tissue compared to glandular tissue [284]

1.4.4 Differentiating normal from cancerous tissue using water in breast cancer

There are a number of techniques that have assessed the differences in water content in breast cancer.

Terahertz pulsed imaging is a non- invasive optical imaging technique that is particularly sensitive to water and lipids (the main areas of spectral interest in HWN RS). In a study using 22 samples of human breast tissue, Terahertz imaging could differentiate between normal and cancerous tissue with good diagnostic accuracy, the increased water content in cancerous tissue hypothesised to be a major discriminatory factor in this [287]. Subsequently a more clinically relevant handheld probe has been developed that had reasonable overall diagnostic accuracy of 75%, with a sensitivity of 86%, but a poor specificity of 66% [288]. This study was limited by the method of measurement, as the probe measured both tumour and normal tissue at the same time, and relied on matching the pixels of measurements with specimen photographs, which could easily be mis-matched after the tissue has been through the specimen processing pathway. Another limitation of this technique was that exact water content changes were not able to be calculated.

A handheld DOS probe has been used in a series of studies, where spectra were taken over a number of points on the skin over tumour containing breast, and the contralateral normal breast. When using average readings over the affected breast vs the contralateral normal breast, they found a significant increase in water concentration of over 50% on the affected side, and decrease in bulk lipids of 20% [276]. As the readings were averaged over the breast which contained normal breast tissue as well it could be assumed that if the specific cancerous were isolated and measured the difference may have been greater. It was also noted that there was significant physiological variation in the readings of healthy tissue between patients. The same group used peak tumour readings compared to average readings from the normal breast, and came to similar conclusions with regard to the changes in water concentration and lipid concentration [289]. However, the sample size was small with 11 patients and the tumour concentration range was considerable and overlapped with the normal tumour range, but still reached statistical significance.

Water concentration in breast tissue displays large physiological variations between patients, therefore using this as a sole measure of malignancy is likely to be inaccurate. It has been seen that areas of tumour not only have an increased water concentration, but also a decreased lipid concentration compared to normal tissue.

The inverse relationship between water and lipid content is seen in studies using MRI for tumour diagnosis. Using Magnetic Resonance Spectroscopy in 15 patients with cancer, and 14 controls with an average age of 47 years, Jaganaathan *et al.* showed that the water to fat ratio increased in cancerous tissue [290]. Spectra from tumour showed significantly higher water to fat ratio (6.0 +/- 6.9) than normal tissue (0.35 +/- 0.26)(+/-SEM). They also demonstrated a degree of variability in tumours. The main aim of the study was to illustrate that water to fat ratio decreased with response to neoadjuvant chemotherapy, and did not investigate the diagnostic ability of the water to fat ratio for cancer diagnosis. In a study using MRS on 68 cancers with healthy controls, the fat fraction was calculated and this measure was used to demonstrate that malignant tissue had significantly lower fat fraction (median 0.12) than normal tissue (median 0.39), concluding that this was primarily due to increased water concentration in cancer. However, the sensitivity of using the fat fraction to diagnose breast cancer was only 75%, probably due to the large ranges of absolute fat and water measurements [291].

DOS has also been used to investigate the fat to water ratio as a discriminatory factor in differentiating normal from cancerous tissue. It is consistent with the MR data, that cancerous tissue has higher water content, and decreased lipid content, compared to normal tissue, both *in vivo* [283] and *ex vivo* [292, 293] .

1.4.5 Differentiating normal from cancerous tissue using water with HWN RS in other cancers

These studies demonstrate that water content or a measure of water fraction is a useful biological feature in differentiating breast cancer from normal breast tissue. HWN RS has been used to quantify the difference in water content between normal and cancerous tissues, mainly in protein rich biological tissue. The spectral peaks of the C-H stretch of protein and lipid at 2935 cm^{-1} and the OH

stretching band of water at 3390cm^{-1}, and the ratio between these peaks, or the ratio between the area under the curve of these peaks has been used to calculate water content in porcine brain tissue [294], the stratum corneum in skin [295] and eye lenses [296].

This technique was used in the assessment of Oral Squamous cell carcinoma with HWN RS to calculate water content [174]. Using freshly excised tongue resection samples, a number of measurements were taken with a HWN RS probe from tumour and normal tissue, before the specimens were fixed and processed for a histopathological diagnosis. They found that water content was 20% higher in cancerous than normal tissue, and that setting the discriminatory value for cancer of water content >69% gave them a sensitivity of 99% and specificity of 92% for differentiating normal and cancerous tissue. The same group used a similar methodology in mandibular bone resection specimens and found a significantly higher water concentration in cancerous tissue, and with a discriminatory water concentration cut off value of 60% had a 99% sensitivity and 83% specificity for detection of cancerous tissue [240]. The water concentration for these studies were calculated based on a calibration using protein/water mixes, whereas the biological tissues that were measured had contribution of lipid signals, so the absolute numbers calculated may not be accurate. The samples were also washed with saline prior to measurement, meaning that fatty (hydrophobic) areas would be unaffected, but possibly more metabolically active, and hydrophilic areas (tumour containing cells) may have taken up water, thus altering and exaggerating the differences in water concentration results seen. Despite these limitations, it still holds true that the changes in the protein to water ratios can be used to differentiate between normal and cancerous tissues. A further limitation of these studies is that Raman microscopy with a 4μm sample area was used to gain spectra, meaning that with the equipment employed in these studies it would be impossible to analyse the entire resection margin of a specimen within a clinically relevant time.

Mo *et al.* obtained measurements from normal and dysplastic cervical tissue with HWN RS [272]. They observed that Raman intensities at the OH band region of 3100-3700 cm^{-1} was significantly higher in dysplastic tissue, indicating that water content was higher in the dysplastic tissue compared to normal cervical tissue,

and this was used in the diagnostic algorithm along with changes in the protein/lipid region of 2800-3050 cm^{-1} to differentiate between normal and dysplastic tissue. However, signal intensity alone may not provide quantification of water changes as they can be affected by other changes in optical properties such as scattering or absorption and water concentration changes alone were not calculated, so the contribution of changes in water concentration to the diagnostic algorithm cannot be assessed.

1.4.6 Differentiating normal from cancerous tissue using water with HWN RS in breast cancer

HWN RS has been used to differentiate between normal and cancerous breast tissue in mammary rat tissue [271]. In this study, measurements were taken *in vivo* transcutaneous, *in vivo* skin removed and *ex vivo* biopsy from 20 rats. The *in vivo* skin removed and *ex vivo* biopsies demonstrated that the intensity band associated with lipid structures (at 2854 cm^{-1} , 2895 and 3010 cm^{-1}) was higher with normal tissue than with cancerous tissue, and the protein band at 2937 cm^{-1} was higher in the cancerous tissue. It was also noted that the water signal at 3100-3500 cm^{-1} was higher in cancer tissue than in normal and benign tissue (Figure 1-20), consistent with there being less lipid and more water in cancerous tissue compared to normal breast tissue. However, the analysis used the intensity band associated with the proteins and lipids and did not include the signal contribution from the water peaks, so the changes in water peak were not statistically analysed. Spectra were analysed by Principle Component Analysis and LDA analyses with cross-validation which demonstrated a sensitivity and specificity of 100% of distinguishing normal from cancerous tissue in the *ex vivo* biopsy tissue. However, the diagnostic role of the differences in water content between the normal and cancerous environments cannot be determined. Another limitation is that HWN RS changes in induced tumours in rats may not be relevant to spontaneous tumours in humans.

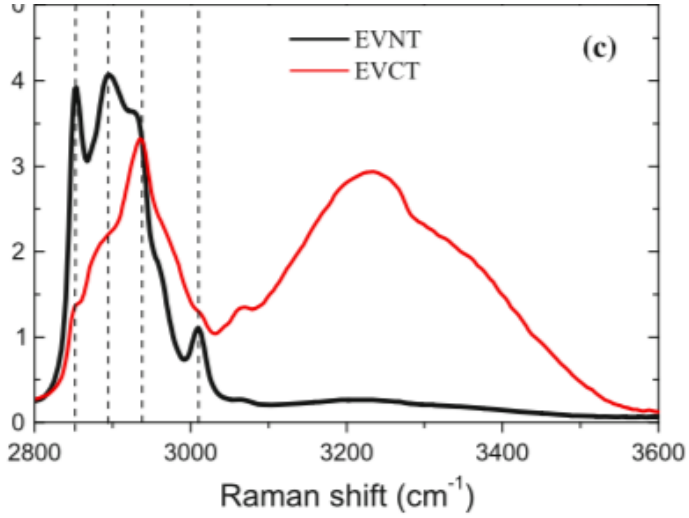


Figure 1-20 Figure of Raman spectra showing the mean spectra of normal and cancer tissue in a selected HWN region for Ex vivo biopsy for normal (EVNT) and cancerous (EVCT) breast tissue from rats.

Reprinted by permission from Springer Nature GmbH: Springer Nature, Theoretical Chemistry Accounts, High-wavenumber FT – Raman spectroscopy for in vivo and ex vivo measurements of breast cancer, A F Garcia- Flores, L Raniero, R A Canevari *et al.*, 2011 [271]

Abramczyk *et al*. aimed to investigate the vibrational features of the OH water stretching bands using Raman Spectroscopy and HWN RS in human breast tissue. Bulk samples of fresh breast tissue were analysed in a controlled humidified environment in a reservoir of aqueous salts. Initial experiments confirmed that in human breast tissue, water was a major discriminatory factor in differentiating normal from cancerous tissue with HWN RS, and that cancerous breast tissue had a higher water content and less lipid than normal breast tissue [297] (Figure 1-21). Subsequently they interpret the vibrational features of water in breast tissue to represent interfacial water, and note that the lack of water in normal breast tissue is likely due to the hydrophobic adipose tissue, which is not present in cancerous tissue [298]. There were limitations to this work as the specimens were selectively sampled, highly pre-processed and measured under controlled laboratory conditions. There was no summary presentation of the results from all measured samples, and no statistical analysis of the diagnostic ability of the HWN RS data to differentiate between normal/abnormal breast tissue, therefore variations between patients or an assessment of the technique for clinical use cannot be done. Therefore, whether these HWN RS findings are found in the *in vivo* or clinical setting with *ex vivo* tissue is yet to be evaluated.

A recent, more clinical, study by Liao *et al.* has used HWN RS in fresh frozen breast tissue to 'screen' specimens of human breast tissue for areas of potential tumour prior to using the FP region for definitive diagnosis [299]. The technique did show promise at highlighting some areas of potential tumour in specimens, however, there were a number of limitations. The background fluorescence of breast tissue was high with the laser excitation they used for obtaining the HWN region (671 nm) which resulted in no spectral information being gained regarding protein or water signals. This meant the 'screening' tool assessed for the presence or absence of lipid only, with no further biochemical information from the HWN region. From the current literature, the difference in water content between tumour and non-tumour tissue may be essential for diagnosis, and this is not assessed in this study. The study was also performed on a, presumably, small number of specimens as no sample size is given, with no statistical assessment of the technique's accuracy in terms of sensitivity or specificity, meaning the diagnostic ability of HWN RS remains unassessed, and the clinical relevance of the technique cannot be determined.

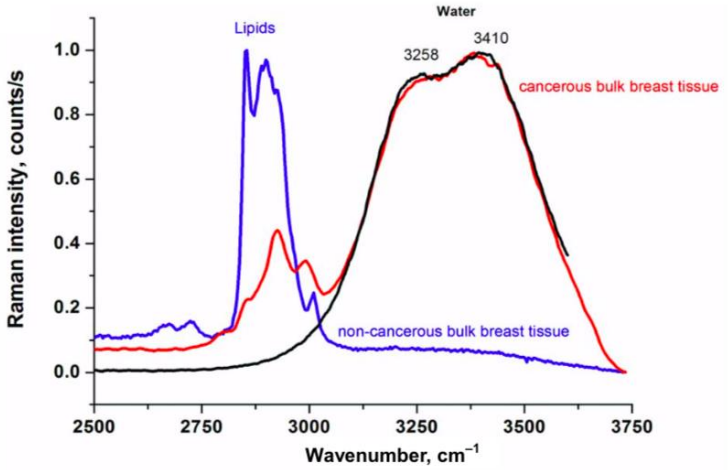


Figure 1-21. HWN Raman spectra of normal (non-cancerous), and cancerous human breast tissue (infiltrating ductral carcinoma) compared to the bulk neat water in bulk tissue.

Note the difference in the signal intensity of the OH stretching vibration of water in cancerous tissue compared to that of normal tissue. Reproduced from reference [297].

1.4.7 Summary

From review of the available literature, cancerous breast tissue has a higher water content than normal breast tissue, and this may be used as a discriminatory

feature in differentiating normal from cancerous breast tissue. However, looking at water alone shows that water content varies between patients, and even within patients depending on hormonally influenced factors. There is an inverse relationship between water and lipids which has been demonstrated in cancerous tissue, that is, that cancerous tissue has a higher water content and lower lipid content compared to normal breast tissue, and this may also be useful for discriminatory diagnosis.

Studies using techniques other than RS utilising the changes in water content or a fat to water ratio for diagnosis have predominantly had a small study sample size and been *in vivo*; specimens that require intraoperative margin analysis are likely to be freshly excised *ex vivo*, and are likely to have different hydration properties due to their lack of blood supply. Studies using the HWN region, looking at water content for diagnosis, have been limited to rats or highly experimental conditions, the protocols being of little relevance to translating this knowledge to the clinical environment.

As a solution to the limitation of fingerprint RS of time taken to analyse a sample, it is proposed that the potentially quicker technique of HWN RS and investigating the potential of water and/or water to fat ratio changes in discriminating normal from cancerous breast tissue may be able to provide intraoperative margin analysis of surgically excised breast specimens.

CHAPTER 2:
Developing a Raman system for intraoperative margin analysis

2.1 Introduction

Although a Raman Spectroscopy (RS) system will share the common basic components as outlined in previous sections, each system is unique and tailored to the aims of the experiment, specimen to be measured, and the environment in which it is being used. The initial stage of the project was to develop a RS system capable of obtaining measurements from breast tissue to elucidate the biochemical features in the High Wavenumber (HWN) spectral region of normal and cancerous breast tissue. These spectral features need to be defined and diagnostic processes validated before the technique can be used for Intraoperative Margin analysis (IMA).

Therefore, the requirements for our RS system were:

- Ability to measure the HWN region Raman spectrum
- Ability to quantify changes in the water concentration and fat/water ratio
- Suitable for use in *ex vivo* breast tissue specimens
- Suitable for use in the clinical environment
- The footprint of the system needed to fit within the space allocated to it in the hospital, where the clinical measurements would be taken

This chapter demonstrates the development of a Raman probe that fulfils these needs. Breast tissue phantoms were used to assess the ability of each potential Raman system to measure the HWN region and obtain accurate water measurements. Potential clinical/logistical problems that may be encountered during measurements were then tested, to ensure the Raman system was capable of obtaining high quality measurements in breast tissue within the clinical environment.

2.2 Materials and Methods

2.2.1 Measured materials

Breast Tissue Phantoms

Tissue phantoms are used extensively in Raman Spectroscopy for the technical development of RS systems and to provide a 'proof of concept' ability of systems to take accurate measurements prior to using that system in a clinical environment. The simplicity of a tissue phantom in a controlled laboratory environment underestimates the difficulty of achieving similar measurements in a complex tissue sample in a clinical environment, however, they are useful to validate basic functionality of a system. A variety of phantoms may be used to assess different aspects of the intended Raman measurements. Previously used tissue phantoms in the field of spectroscopic breast cancer diagnostics include porcine tissue [300], chicken breast [301], soybean oil and water [302], a solidifying agent TX151 with water, oil and surfactant [303], pure gelatine and water phantoms [304], and gelatine, water and intralipid phantoms [278]. A single simple phantom cannot be constructed that replicates the exact spectroscopic properties of the real tissue specimen. The phantoms used in this chapter are Gelatine, Gelatine/intralipid constructed phantoms, and pork meat.

The initial challenge is to identify the tissue composition and physiological concentrations the tissue phantom is attempting to mimic. Breast tissue is primarily composed of epithelial tissue and stroma (protein), adipose tissue (fat) and water [284]. From an MRI study (with a small sample size of n=8) it was predicted the breast was composed of 67% fatty tissue and 33% parenchymal tissue [305], however this was with a wide degree of inter-subject variation. A more recent MRI study of 306 women suggests the range of fatty tissue is from a mean of 53% in women with a mean age of 50.9 to 71% in women with a mean age of 20.8. The water content of normal breast tissue has been estimated to be around 27% with Diffuse Optical Spectroscopy [306], 40 – 60% with NIR spectroscopy [281], and around 20 – 80% with MRI [280] demonstrating a very wide range in the composition of breast tissue. Breast composition varies significantly between women in terms of mammographic density (ranging from fatty breast to dense and homogenous [307]) and water content [284], and within a women's lifetime due to hormonal factors such as pre/post menopause, Body Mass Index (BMI), time of menstrual cycle and exposure to the combined oral contraceptive pill [280, 282, 284].

Due to these large variations in absolute differences, looking at the change in ratios between breast tissue constituents has been used for diagnostic purposes. Previous work using MRI and Diffuse Optical Spectroscopy has demonstrated that the differing ratios between water and fat relate to the differences observed in breast tissue density[280], and breast tissue composition changes were observed with increasing age, menopausal status and Body Mass Index (BMI) [284]. Previous work in Raman Spectroscopy using the fingerprint region has demonstrated that the differing ratio between collagen (protein) and fat differentiated between normal, fibrocystic changes and cancerous breast tissue [180, 262]. Ship *et al.* identified major spectral features in normal breast tissue as corresponding to collagen and other proteins in stromal tissue and lipids in fatty tissue [189].

This suggests that protein, fat and water are the major constituents in breast tissue that can be examined for determining differences in breast tissue composition and disease (such as cancer) that can be readily measured with Raman Spectroscopy. These constituents are the primary biochemical signals that can be detected with HWN RS [105]. Therefore, the breast tissue phantoms were composed of these constituents, with the ability for the ratio between them to be changed.

Justification of type of Breast Tissue Phantoms used

Gelatine based tissue phantoms are attractive in preliminary work due to their simplicity and the ability to set solid in the shape of the mould, which means that any physical logistical aspect of measurements could also be assessed. Constructing phantoms are preferable for initial experiments compared to meat-based phantoms because they are homogenous, the phantom constituents are known and the percentage water concentration and the phantom constituents is a known entity. This is a necessary attribute for the tissue phantom in measuring the ability of the RS system to measure in the HWN region and the ability to quantify changes in water concentration. A previous study by Masson *et al.* demonstrated the ability of gelatine phantoms to reproduce physiological levels of water concentrations, the HWN Raman spectral profile of the phantoms and the ability to accurately detect changes in water concentration [304] (Figure 2-1).

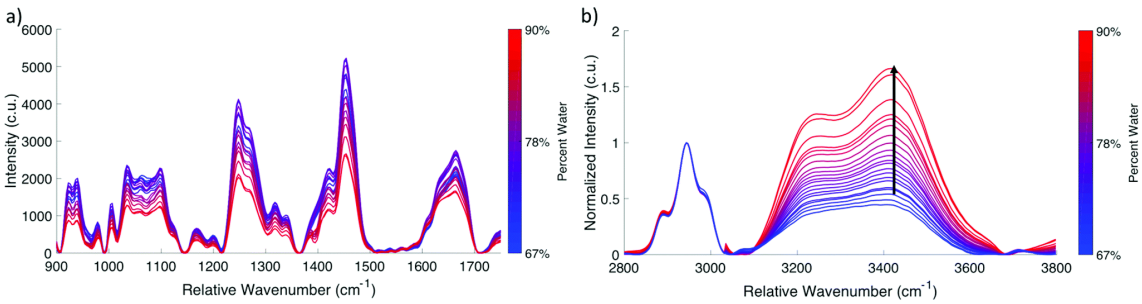


Figure 2-1 HWN Raman spectra from a set of gelatine phantoms of different water concentrations. Spectra were normalised to the protein peak at 2940cm^{-1}. Reproduced from reference [304] Reproduced by permission of The Royal Society of Chemistry

Gelatine based phantoms can also be constructed to be more complex, and set with the addition of a lipid (such as Intralipid [278, 308]), and scattering agents such as india ink[308] or nigrosin [278]. Prior knowledge of the optical characteristics and properties of these phantom constituents such as absorption [309] and scattering [308] as well as the biochemical properties such as the changes in bound versus free water [278] with gelatine concentration makes them a well understood and reliable tissue phantom.

There are some limitations of intralipid /gelatine based phantoms. The lipid concentration cannot reach the physiological concentrations of the high fat concentrations found in breast tissue, as intralipid is only 20% soyabean oil (fat). As an oil and predominately unsaturated fat that is liquid at body temperature, it may have very different characteristics to that of the complex fat composition of human adipose tissue that is solid at body temperature [310]. The optical scattering properties of intralipid may affect the Raman cross section of phantoms with a change in intralipid concentration which could make it difficult to ascertain if it is changes in fat content or scattering properties that are being measured [311]. However, the simplicity in synthesising these phantoms makes them highly reproducible and means previous findings are applicable to our own phantoms, along with the ability to manufacture a variety of varying lipid/protein/water concentrations means they are the optimal phantom to be used for preliminary work with a Raman system investigating the HWN region.

2.2.1.1 Method for Gelatine based Breast Tissue Phantom synthesis

Pure Gelatine Phantom

Different concentrations of gelatine phantoms were made by mixing gelatine powder from Bovine skin (Sigma Aldrich,Germany) with distilled water to a final total weight of 10 grams in a water bath at 55°C and stirred with a magnetic stirring bar until completely dissolved (20 – 30 minutes) then poured into moulds and cooled at 5 – 7 °C whilst covered to avoid evaporation for 12 – 24 hours before measurement [312]. Water concentrations achieved were 85%, 87.5%, 90%, 92.5% and 95% (Figure 2-2). These water concentrations were chosen as the phantoms would not set with water concentrations of > 95%, and when water concentrations of <82.5% were attempted the high concentration of gelatine led to saturation of solution. Samples were kept in petri dishes with lids between measurements to prevent dehydration.

95% 92.5% 90% 87.5% 85%

Figure 2-2 Photograph of the pure gelatine phantoms at decreasing water concentration

Complex intralipid/gelatine Phantoms

The same method was used for the construction of the complex intralipid/gelatine phantoms. With a final weight of 10 grams, a fixed weight of gelatine (1 gram) and india ink (2μl)(Fischer Scientific,Pittsburgh,USA) were used in varying amounts of distilled water and Intralipid 20% (Fresnius,Frankfurt,Germany) with final compositions and predicted fat and water % shown in Table 2-1. Figure 2-3 shows the complex phantoms. Intralipid is composed of 20% soyabean oil, which is 100% lipid, and 3.45% other constituents (egg yolk and glycerine) [313], and the remainder is water allowing for the calculation of the total fat concentration and water concentration.

In order to produce complex phantoms analogous to breast tissue the scattering and absorption qualities had to be considered. Intralipid is a scattering fat emulsion, the scattering coefficient changing with the intralipid concentration[311]. As the priority with these tissue phantoms was to provide a varying range of fat

concentrations, a range of intralipid concentrations (including high concentrations of 90%) were measured rather than attempting to reproduce physiological levels of scattering [314]. India ink was used to provide an absorber, concentrations in which (0.2μl per ml) had been previously used in similar work with intralipid phantoms [308]. As these were not deep Raman measurements, the importance of physiological scattering and absorption coefficients are less, and are unlikely to affect the results, however they are considered as deep Raman measurements may have been taken later in the project.

NAME OF PHANTOM	INTRALIPID (ML)	WATER (ML)	GELATINE (GRAMS)	TOTAL FAT (%)	OTHER CONSTITUENTS (%)	CALCULATED WATER CONCENTRATION (%)
IL 10	1	8	1	2	0.345	88
IL 30	3	6	1	6	1.035	83
IL 50	5	4	1	10	1.725	78
IL 70	7	2	1	14	2.415	73
IL 90	9	0	1	18	3.105	68

Table 2-1 Table showing the constituent parts of the complex intralipid/gelatine phantoms

IL 10 Il 30 IL 50 IL 70 IL 90

Figure 2-3 Photograph of complex gelatine phantoms with increasing concentrations of intralipid

Bovine Serum Albumin (BSA)

BSA and water solutions were made to provide a standard against which to test the performance of gelatine phantoms. BSA (Sigma Aldrich,Germany) and distilled water solutions were made to a total weight of 1 gram to make solutions of final water concentrations of 95, 90, 85, 80, 75 and 70% by weight.

2.2.1.2 Recreating the challenges of the surgical environment

Introduction

After establishing the ability of the Raman system to accurately measure the HWN region, the suitability for use in measuring fresh breast specimens with a view to IMA needs to be evaluated. Other optical techniques which have shown

scientific promise have had issues with measuring fresh samples in the clinical environment, such as optical coherence tomography that could not gain a signal in specimens that had been contaminated with blood or encountered cauterised tissue [145], and previous work performed by the Stone group showed that blue dye can affect Raman signal [315]. As these challenges would be encountered in any clinical application of a Raman system for IMA, it was imperative we designed a system that could function in the presence of these potential signal altering dyes.

Pork Tissue

Porcine tissue has been used extensively in previous Raman studies, as it contains fat and muscle (protein) and thus grossly replicates the chemical composition of human breast tissue [316]. The obvious visible distinction between, and the existence of areas that are almost exclusively, fat and proteinaceous muscle tissue means these areas can be measured to assess spectroscopic characteristics in these different tissue types. The benefit of using meat phantoms is to assess possible 'real world' clinical issues in obtaining measurements from breast tissue, such as the effect of surgical dye [315], diathermy [317] and blood / haemoglobin contamination[292] which is not possible with gelatine phantoms.

Pork chops with distinctive meat and fat portions were selected from local supermarkets for measurements. They were measured fresh on day of purchase, or frozen on day of purchase at -80 °C and thawed for 24 hours in a refrigerator at 4 – 8 °C prior to measurements. Samples were kept in petri dishes with lids between measurements to prevent dehydration.

Patent Blue Dye

Patent blue dye was used to evaluate the ability to obtain Raman measurements in the HWN region in tissue that had been stained with Blue dye. Blue dye is used in the dual localisation technique as recommended by NICE for performing sentinel lymph node biopsy [56]. Blue dye is injected and travels within the lymphatic system to identify the first or 'sentinel' lymph node that drains the breast undergoing surgical excision. The theory is that if the cancer has spread into the lymphatics, it will be to these sentinel nodes first – thus if the sentinel nodes are negative, it is unlikely any other axillary lymph nodes have metastatic cancer [84]. Blue dye is used in all sentinel lymph node operations which are carried out within

the breast cancer excision operation. Blue dye is injected under the nipple subdermally, and thus frequently stains the breast tissue that is excised, with the potential to affect Raman measurements. Patent Blue V sodium salt (Sigma Aldrich, Germany) was dissolved in distilled water to achieve a 1% concentration and was applied to pork tissue. Methylene blue and patent blue dye are both used in sentinel lymph node operations, at concentrations of 1 – 2.5%[318], the concentration of staining on the actual breast tissue varies widely according to site of tumour, site of injection, length of operation, handling of specimen so the amount of staining on the tissue was based on a judgement of whether the amount looked similar to the degree of staining during breast operations (Figure 2-4).

Figure 2-4 Photograph of pork fat (left) and meat (right) that have been stained with 1% patent blue dye.

These specimens are as heavily stained as breast tissue specimens can be during a breast operation.

Blue dye Gelatine phantoms

To further investigate and quantify the effect of blue dye on the ability of Raman systems to obtain Raman spectra, blue dye gelatine phantoms were produced. 90% water gelatine phantoms were produced as previously described, with blue dye pipetted into the gelatine/water mix prior to stirring and cooling at a final concentration of 0.01% or 0.1%.

Diathermy

During breast operations, electrocautery diathermy is used in the excision of breast tissue. This can lead to burning / carbonisation of the breast tissue, and on histological examination of the tissue can cause significant ‘diathermy artefact’. This is where diathermy limits the assessment of the margins of the excision due to carbonisation and distortion of cellular architecture[319]. Therefore, this problem is not limited to optical techniques and negatively impacts the ‘gold standard’ of histopathology, however, it is important to recognise its effect on Raman spectra so it can be accounted for. A soldering iron was used to burn /

carbonise edges of pork tissue to replicate this, and assess if there was any effect of burnt /carbon on obtaining HWN Raman measurements (Figure 2-5). Pork fat simply melted when using a soldering iron – which does occur during breast operations, so only carbonised pork meat was measured.

Figure 2-5 Pork meat that has been burnt with some carbonisation evident

Haemoglobin

Blood is another pigmented substance that is frequently present on surgically excised specimens which may be Raman active and interfere with obtaining Raman measurements.

Haemoglobin is the major pigmented constituent of blood, and is known to have be Raman active in the fingerprint region, and to a lesser extent in the HWN region [320], however work has mainly been with visible light laser such as 532nm excitation. Normal adult haemoglobin concentrations in venous blood are 13-17 %, however, the concentration on a surgically excised specimen is likely to be less as whole blood is mixed with melted fat, blue dye and tissue fluid. To assess the potential for blood contamination of surgically excised specimens, haemoglobin (Hb from bovine; Sigma Aldrich,Germany) was dissolved in distilled water to achieve concentrations of 7.5%, and the porcine tissue soaked in the haemoglobin solution until measurement (Figure 2-6).

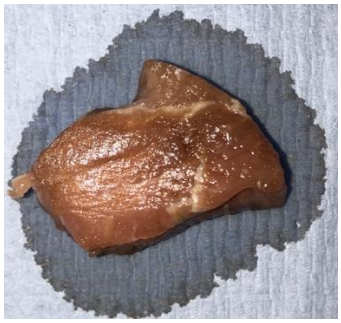


Figure 2-6 Pork meat stained with Haemoglobin solution

Specimen coverings

In order to take repeated measurements, and measurements with the same probe on a number of different samples, the probe which is interacting with the

specimen must be able to be cleaned or disposable, to prevent cross-contamination or tumour seeding leading to inaccurate histopathology results. The Raman needle probe used for measurements has the fibre optic needle recessed within a hypodermic 22G needle. Theoretically this means the probe is capable of having a disposable needle tip, or a needle that can be cleaned while not disturbing the optical components. It is also possible to construct a single – use needle probe cover which would also ensure cleanliness. However, for experiments taking surface measurements from breast specimens, this would introduce unnecessary time delays and costs to each measurement. A simple, disposable barrier such as Clingfilm between the needle tip and the specimen to be measured would ensure no cross-contamination and be more cost-effective. Measurements were taken of pork meat and fat with and without a covering of clingfilm to investigate if this affected the Raman signal.

2.2.2 Raman Spectroscopy equipment configuration and spectral acquisition

2.2.2.1 Raman Microspectroscopy

Preliminary measurements on gelatine phantoms to confirm method of data analysis for future Raman systems, and to validate gelatine phantom construction reliability were taken on a Renishaw InVia confocal spectrometer system (Renishaw, UK). Measurements were taken using a NPlan objective (Leica, Germany) with magnification X50. An excitation source of a 785nm laser was used with an output of 300 mW to collect spectra in the wavelength range of 2100 – 2750 cm^{-1}. Calibration of the system was performed daily with silicon, green glass, PTFE and Neon Argon lamp. Point measurements were acquired with an exposure time of 5 seconds with 12 accumulations.

2.2.2.2 Needle Raman probe

For all other Raman measurements, a Raman needle probe was used. It has been developed, and is similar to, a previously described needle probe [164]. The purpose of such a probe is that, as it fits within a standard hypodermic needle, it can be used to probe within biological tissues, with a view to having *in vivo* diagnostic use and real time subcutaneous pathological diagnosis. This probe was used at this stage of the project as it can provide measurements from small and specific areas of specimens which could then be subsequently matched with

the histopathology, allowing an accurate diagnosis. With this approach, spectra can be gained from normal and cancerous tissue.

The needle probe is not suitable for performing IMA, however it allows this preliminary work to be carried out. Once the diagnostic algorithm has been validated, a probe that is suitable for analysing the resected surface of a breast tissue specimen can be developed. We envisage this would use the same experimental set up with a different probe for light delivery and collection more suited to assessing a large surface. The basic ability to differentiate between normal and cancerous tissue using HWN RS, which is this project's aim, would underpin the ability of this probe to provide IMA.

The needle probe is pictured in Figure 2-7, it is a standard stainless steel 22 gauge hypodermic needle which provides protection to the fibres, and allows cleaning between specimen measurements. Within this probe is a bundle of low-OH silica optical fibres stripped of coating (0.22NA, Thorlabs), arranged as 6 collection fibres around 1 excitation fibre. This gives an area of collection of 5.19 x 10^4 µm^2.

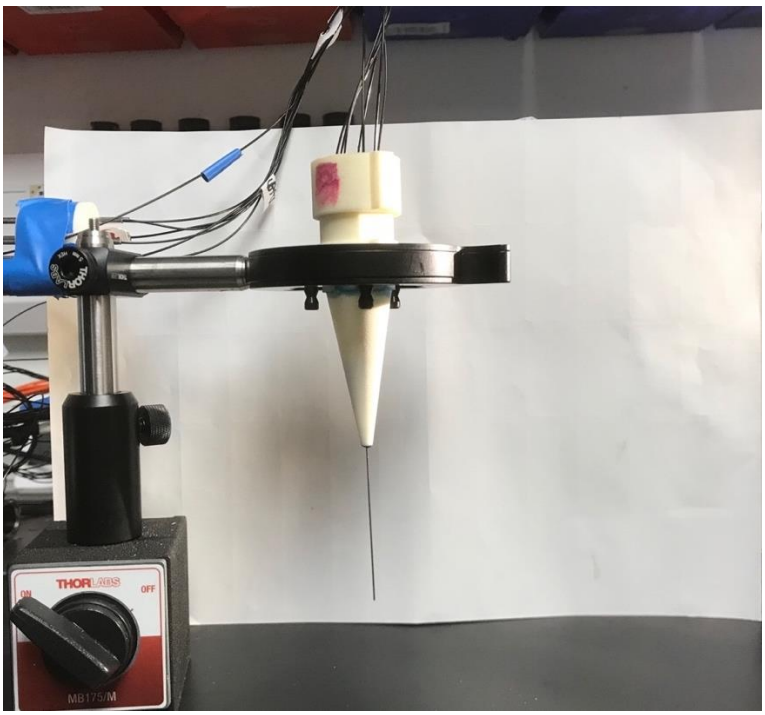


Figure 2-7 A picture of the needle component of the Raman needle probe which delivers and collects light to the specimen for Raman analysis.

Covered fibres can be seen entering the unit at the tip, where they are then uncovered and pass down the needle to tip.

Over the course of the project, some of the collection fibres became dirty or broke, giving poor signal, the data from these fibres was omitted from analysis. The number of working fibres ranged from 1 – 6.

The exact equipment used is specified in each section for each potential Raman system, but all were configured in the same basic way as depicted in Figure 2-8. Excitation was delivered at either 680nm or 785nm delivered by an IPS spectrum stabilised laser module (Innovative Photonic Solutions,USA) through either a 680nm or 785nm laser clean up filter (Thorlabs,NJ,USA) and 830nm edge filter to the needle tip, providing a maximum power of 100 mW at 680nm or 250mW at 785nm to the sample. The light was then collected through the collection fibres through to the entry port on the spectrometer - Kaiser Holospec Imaging Spectrograph (Kaiser optical systems inc, Ann Arbour,USA) with a broad grating which is coupled to a camera – either a deep-depletion Charge – Coupled Device (CCD) camera (PIXIS 400 BRX, Princeton Instruments, NJ, USA) thermoelectrically cooled to -80°C or an InGaAs camera (iDus InGaAs 1.7μm, Andor, Belfast, UK) cooled to -85°C using water cooling (Julabo, Germany).

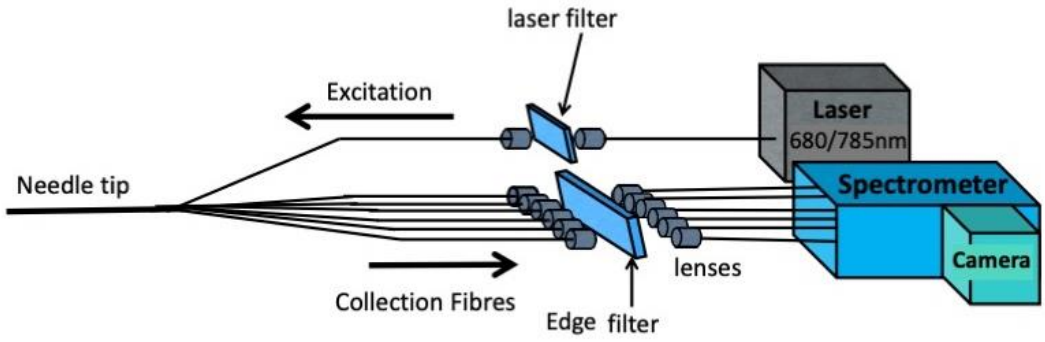


Figure 2-8 Schematic diagram showing the Raman needle probe set up, modified from ref [164].

Sample measurements

Breast tissue phantoms were placed on a PTFE block (which has no discernible spectra in the HWN region) and measured with the tip of the Raman needle probe in contact with the specimen. The acquisition times and number of accumulations varied according to the Raman system, and are detailed in the corresponding sections. All phantoms were measured in three different places to provide a

representative result for the whole phantom, and to ensure any heterogeneity in the phantom constituent concentrations was identified.

2.2.3 Calibration

Calibration in Raman spectroscopy is the process that is used to correlate the observed spectral frequencies to their true values. Calibration is essential in performing Raman spectroscopy because no two Raman instruments measuring the same sample would give identical raw spectrum without calibration.

The need for calibration is due to the large number of sources of variation between measurements, such as temperature, and between instruments, such as alignment of optical systems, all of which effect the raw spectrum. These variations can cause errors leading to miscalibration of the instrument, and discordant results. Within the literature, the Raman shift values that are reported and taken as standards, can vary significantly; for the example of ethanol in the HWN region there was a difference of 2-5 cm^{-1} between different labs [321, 322], which is not uncommon [323].

Common sources of variation are[324]:

- Rotation of spectrograph diffraction grating or CCD
    - Spectrographs allow rotation to allow reading of different regions of the spectrum, however a misaligned grating by even a small amount is a potential source of error in calculating wavelength.
- Displacement of the camera
    - An inadvertent change to the camera position (i.e. when replacing) can cause a shift in camera pixels.
- Changes in laser excitation wavelength
    - Although stable lasers aim to achieve a single coherent wavelength, this can change over time, and can have an appreciable effect on wavenumber calculation
    - If using reference spectra that were gained from a laser with a different wavelength to the one being used, it can affect the peaks due to resonance effects[323].

- Temperature
    - Not only is Raman scattering dependent on temperature, but molecular structure is also influenced by temperature.
    - The ambient temperature has a significant influence on the camera's detection of Raman scattering. At room temperature there is significant thermally generated background noise which can reduce the signal to noise ratio and the ability to detect Raman scattering, and so deep cooling of the camera is necessary to reduce this 'Dark noise'.

Therefore, calibration is performed to ensure that there is internal validation – that the same instrument will give the same result on different days, and external validation – that different instruments in different places will give the same result. This is essential if Raman spectra results are to be reproducible, interpretable and informative.

Frequency Calibration with Raman shift standards

There is no universally standardised method of calibration for Raman instrumentation, and so there are a number of methods to calibrate an instrument. The most common method is to use wavenumber calibration based on known frequency standards such as neon-argon lamps, as these standards have a well-documented and wide range of spectral atomic emission peaks[325]. Another calibration method is to use known frequency Raman shift standards. In this method a substrate is used that has well characterised and documented peaks that have an assigned Raman shift number, the substrate is measured under the same conditions that the experimental samples will be measured under, and the 'true' peak values of the substrate are mathematically fitted (by a polynomial fit) to the obtained pixel peak values [323]. In this way the instrument's Raman shift axis can be calibrated. The benefits of this method are the laser frequency does not need to be known, as long as it is constant, the sample position does not influence the result, and it is easy to implement [326]. In these experiments, a number of different laser wavelengths and lasers were used, and so it was practical to use this method for calibration, to avoid repeated measuring of the laser frequency.

Finding a Standard

It is suggested that a wide number of at least 10 peaks with a broad range across the interested region are used to optimise the accuracy of calibration across the spectrum [323]. The HWN region has fewer spectral features for shift standards to be calibrated against compared to the fingerprint region, and so a number of shift standards can be used at the same time in order to increase the range and accuracy of calibration [324]. The American Society for Testing and Materials (ASTM) selected a number of materials that could be used as a common Raman shift standard, and examined them in a reproducible manner in a number of laboratories to provide the ASTM standard. Commonly used ASTM calibration standards that have peaks in the HWN region are 4-Acetamidophenol (paracetamol) and acetonitrile [323, 324, 327], and so these were used for our calibration. Additionally, it was found that ethanol produced a clearly defined spectrum in the HWN region, which remained easily distinguishable in the HWN region when using a 785nm excitation wavelength (where the paracetamol spectrum peaks became less distinct). To ensure there were enough spectral peaks in the HWN region ethanol was also used as a calibration standard [321].

The Raman system to be used for the measurements therefore was calibrated using the frequency calibration method using the Raman shift standards of paracetamol, acetylnitrile and ethanol.

Methods of calibration for the Raman Needle probe

Calibration was performed daily, or with any change of equipment. Using the Raman needle probe system, measurements were taken of the calibration standards for the relevant wavelength of either 680nm or 785nm. The calibration standards were paracetamol (Tesco,UK), acetylnitrile (Sigma Aldrich, Germany), and ethanol (Fischer Scientific,Pittsburgh,USA). A single acquisition was taken, with an exposure of 0.5 seconds. The pixel number of the characteristic peaks for each standard were identified and correlated to the ‘true’ Raman shift number from the literature (Table 2-2). These were combined and entered in a table in Microsoft Excel, and regression analysis performed to report the adjusted R Squared and standard error of regression (as a description of fit) and to report the y intercept and 1st, 2nd and 3rd coefficient. These were then used to fit the pixel number to the Raman shift number using a 3rd order polynomial from Equation 2-1:

$$Raman\ shift_p = I + C_1 p + C_2 p^2 + C_3 p^3$$

Equation 2-1 Equation adapted from reference [328]

Where $\lambda_p = Raman\ shift\ of\ pixel\ p$, I = Raman shift of pixel 0, C_1 is the first coefficient (cm^{-1}/pixel), C_2 is the second coefficient (cm^{-1}/pixel), C_3 is the third coefficient (cm^{-1}/pixel).

The degree of polynomial to use is a subject that has been debated, as the user needs to avoid using too low a polynomial and underfitting, or using a high polynomial and overfitting the data. A third order polynomial was used for fitting as it is thought to be optimal for shift accuracy when compared to other order polynomials for pixel fitting, and is commonly used [329, 330]. This produced the calibrated Raman shift axis for the experiments produced at that laser excitation wavelength for that particular experimental session.

CALIBRATION STANDARD	TRUE RAMAN SHIFT CM^{-1}	LASER WAVELENGTH UTILISED
PARACETAMOL	1278	785nm
PARACETAMOL	1323	785nm
PARACETAMOL	1371	785nm
ACETONITRILE	1374	785nm
ETHANOL	1463	785nm
PARACETAMOL	1515	785nm
PARACETAMOL	1561	785nm
PARACETAMOL	1648	785nm
ACETONITRILE	2253	785nm
ACETONITRILE	2293	785nm
ETHANOL	2887	785nm + 680nm
PARACETAMOL	2931	785nm + 680nm
ETHANOL	2937	785nm + 680nm
ACETONITRILE	2943	785nm + 680nm
ETHANOL	2983	785nm + 680nm
PARACETAMOL	3064	785nm + 680nm
PARACETAMOL	3102	680nm
PARACETAMOL	3326	680nm

Table 2-2 Table demonstrating the Raman shift references used for calibration of the Raman shift axis, and the laser wavelength for which they were utilised.

The reference spectra were paracetamol, acetonitrile (both reference [331]) and ethanol [321].

2.2.4 Data recording and Analysis

Raman spectra were recorded using LightField software (Princeton Instruments, USA) when using a Princeton Instruments camera, or SOLIS software (Andor,UK) when using an Andor camera. Raw data was exported and analysed using software Matlab R2019a(The MathWorks Inc, USA). Bar graphs were constructed in Microsoft Excel.

Pre-processing

The calibration protocol for the Raman Needle probe is described in the previous section, which allowed the x-axis of Raman shift to be correctly aligned.

As previously mentioned, the number of working fibres within the Raman needle probe varied between 1 and 6 and only spectra from working fibres was included. Within each fibre, accumulations were summed together, and then summed to the output from the other fibres within that measurement to create a single spectra for each measurement. All data were baseline subtracted to remove background fluorescence using a 1^{st} order polynomial using the lowest count at the beginning of the HWN region, and the end of the HWN region. As we were interested in the protein/fat to water ratio, for visualisation purposes, spectra were then normalised to the CH_2 peak at 2935 cm^{-1}, and the average and standard deviation of the triplicate readings calculated. This allowed visual comparison in the water spectra to be made between samples of varying water content.

Calculation of water content

A central idea to this project is the differentiation between normal and cancerous tissue using water content. As outlined in Chapter 1, previous studies demonstrate that breast cancer tissue has more water than normal breast tissue, and that changes in the fat/water ratio can be used to differentiate between normal and cancerous breast tissue [290, 291]. It would be useful if we could not only identify changes in water content using Raman spectroscopy, but quantify what those changes are, and provide estimates of water concentration. This would allow any results to be compared with results gained from other spectroscopic methods such as DOS [289], and compared with other Raman studies investigating water concentration changes in other disease states [174]. A peak assignation table for the protein /fat CH stretch region is displayed in Table 2-3 and for the water OH stretch region is displayed in Table 2-4.

TYPE OF TISSUE	APPROXIMATE PEAK LOCATION	MOLECULAR INTERACTION	REFERENCE
RAT BREAST TISSUE	2817-2849	CH_2 symmetric stretch of lipids	[103], [271]
HUMAN /RAT BREAST TISSUE	2854	CH_2 (breast)	[297], [271]
	2840 - 2875	CH_3 symmetric stretch of lipids	[103], [271]
HUMAN /RAT BREAST TISSUE	2888-2895	CH3 Lipid (breast)	[297], [271]
COLLAGEN TYPE I	2928-2945	CH_3 Protein (Collagen Tpe 1)	[332], [333]
HUMAN SKIN	2910-2966	CH_3 Protein (Human Skin)	[295]
HUMAN SKIN	3329	NH	[304], [333]

Table 2-3 Table demonstrating peak assignation for the HWN Raman spectroscopy CH stretch region in a variety of human tissues

APPROXIMATE PEAK LOCATION	HYDROGEN BONDING SCHEME	MOLECULAR INTERACTION	REFERENCE
3075	DAA	Fully hydrogen bound	[304]
3200-3250	O-H symmetric stretch	Bound water	[334], [297]
3245	DDAA	Fully hydrogen bound	[304]
3400 - 3450	OH asymmetric stretch	Unbound water	[334], [297]
3420	DA	Partially hydrogen bound	[304]
3550	DDA	Partially hydrogen bound	[304]
3650	-	Free water	[304]
3600-3650	O-H non hydrogen bonded	Free Water	[334]

Table 2-4 Peak assignation table for the HWN Raman spectroscopy OH stretch region

The spectral intensity of a Raman peak is often proportional to the change in concentration of the molecule assigned to that peak i.e. if a substance has more of a molecule then that assigned band will have a higher spectral intensity than a substance that has less of that molecule, and this can be used for diagnostic purposes [232]. We could presume that (using BSA/water mixtures analysed with the InVia spectroscope as per the protocol in 2.2.1.1 as an example) the peak height of the OH stretch of water at 3390 cm^{-1} should change proportionally to the water concentration (measuring the peak height relative to the baseline, as previously described in reference [232]). Figure 2-9 demonstrates that this is not the case, and the maximum peak intensity for the OH stretch region does not vary significantly (one-way ANOVA comparison of means, $P>0.05$) with a changing water concentration. This suggests that another method is needed for quantifying changes in water concentration.

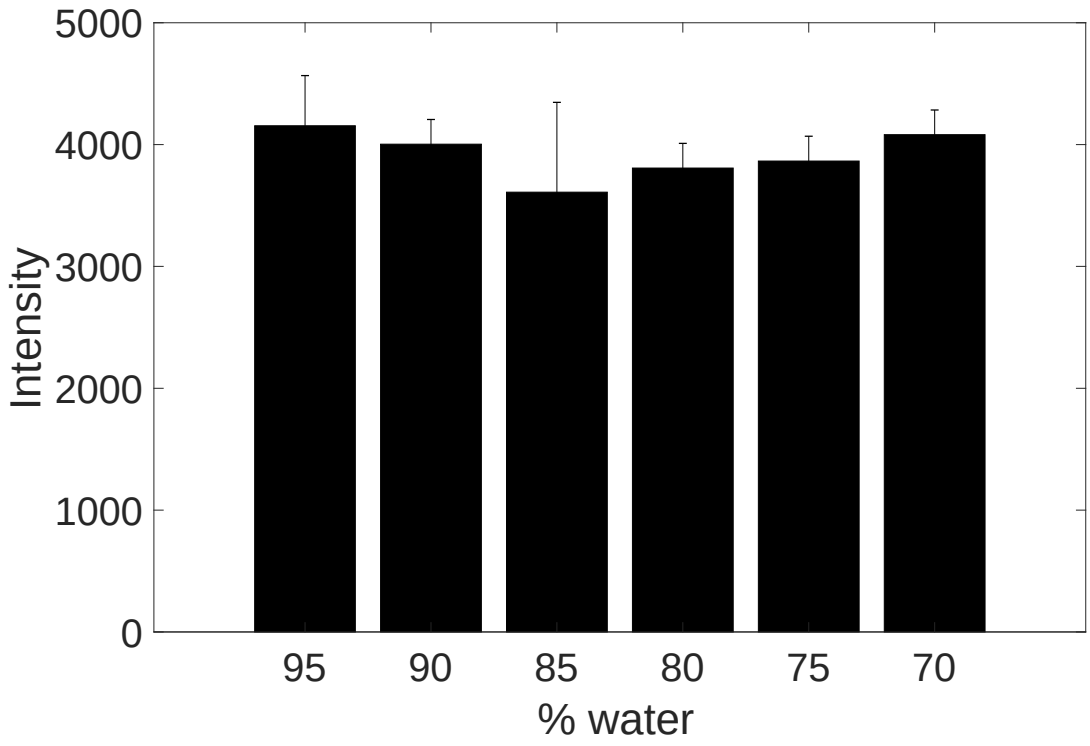


Figure 2-9 Bar chart of maximum Raman spectral intensity at 3390 cm^{-1} at different water concentrations calculated by BSA/water mixtures.

Bars are average of triplicate readings +/- SD.

The majority of other studies quantifying water content changes with HWN Raman spectroscopy have been studying protein rich tissues. Huizinga *et al.* used the ratio between the Raman intensity at the CH stretch of protein at 2935

cm^{-1} and the OH stretch of water at 3390 cm^{-1} in a range of protein/water solutions (including BSA) at similar concentrations of 10-35% water to calibrate and quantify the water concentration in rabbit eye lenses [296]. A similar method using spectral intensity was used by Bauer *et al.* which used traditional methods of drying and lyophilization to confirm the accuracy of this technique to within 3% accuracy [335]. Both of these techniques are dependent on a single intensity reading at the protein peak, which although may be present in fat-laden breast tissue, the relationship between this sharp peak, and the broad multiple peaks of lipid in breast tissue (which is between 2854 – 2940 cm^{-1} [336]) is unlikely to be so simple or linear. Caspers *et al.* used an area under the spectral interval method to calculate the water content in skin, and went on to use it in their work on oral SCC [228, 295]. They calculated the ratio of the area under the curve of OH stretching (3350 – 3550 cm^{-1}) and the area under the curve of CH_3 stretching (2910 – 2965cm^{-1}) and used BSA solutions to find the proportionality constant that related the ratios before applying the method to skin specimens. Although this area-under-the-curve method is attractive as it could be broadened to include lipid regions, the overlap between the protein (2910-2965 cm^{-1}) and lipid regions (2854-2940 cm^{-1}) in the HWN region mean distinction between measuring protein or lipids would be arbitrary. It also uses a specific assignment of the OH stretch of water, whereas water has a much broader range of Raman activity from 3189 – 3770 cm^{-1} [336], and so changes in water spectrum might not be truly reflected by choosing a narrow part of the spectrum.

A more complex model has been suggested using a training set of brain tissue, which has some lipid content. Raman spectra were measured in the training set of wet tissue which was then dried and repeat measurements taken, a PLS model was then applied and achieved good diagnostic accuracy on the test set [294]. However, this method requires access to tissue which would not then undergo pathological analysis for the training set, as it would undergo severe dehydration which may affect diagnosis. This limits its practical use, where this excess tissue may be difficult to obtain.

A method based on area under the spectral curve, whilst avoiding using specific spectral bands was proposed by Masson *et al.* [304]. In this study they were using the HWN region to study water concentration changes in cervical remodelling,

where they used the ratio of the area under the water curve (3035 – 3680 cm^{-1}) to the under the area under the entire HWN region (2850 – 3680 cm^{-1}) to predict water concentration and validated it against wet and dry weights. The reason for using these areas was because cervix remodelling includes significant changes in lipids, and this was captured using this approach, suggesting its suitability for fatty breast tissue. This method also uses all the information from the water peak in its calculation, thus can take into account the varied contributions to the HWN Raman spectra different states of water give. Although this method was validated in protein only (with gelatine/water phantoms), it was subsequently used in lipid containing biological tissue to observe changes in water content. As this method is the only validated method that takes into account the lipid signal, and takes into account the entire water spectrum, this method was used in this project for calculation of water content.

Method of water calculation

In the preliminary breast phantom studies, the relationship between the water/total area ratio and known water concentration was investigated. The area under the curve of the water curve (3035 – 3680 cm^{-1}) and the area under the curve of the entire HWN region (2850 – 3680 cm^{-1}) was calculated, and the water/total area ratio calculated. This ratio was then plotted against the known water concentration (as a ratio) of the breast phantom on the x axis, a 1^{st} order polynomial (line of best fit) was then calculated to assess whether the relationship was linear. This line was also used to calculate the gradient and the y intercept, and the Root Mean Square Error (RMSE) as an indicator of fit.

Statistical analysis

Data is displayed as mean water/total area ratio with error bars +/- 1 standard deviation (SD). Comparison of mean water/total area ratios was performed using one-way ANOVA test. Statistical significance was set to $P<0.05$.

2.3 Evaluating Breast phantoms and data analysis procedures

The InVia microspectroscope was used to assess gelatine phantoms to test the method of data analysis and to evaluate gelatine phantom reproducibility and reliability. The InVia was used for this as it is a fixed, closed, laboratory system which has a defined calibration protocol. Therefore, it allows for easier comparison between experiments, and is a validated piece of equipment. It means any unexpected results or variations can be assumed to be in the phantoms measured, whereas in an unvalidated, open system which is more dynamic such as the needle probe, unexpected results could be due to system set up / calibration rather than the phantoms.

Evaluation of data analysis protocol for measuring water concentration

As already outlined in the previous section, raw water peak intensity does not vary between different concentrations (Figure 2-9 and Figure 2-10 A). However, when the data is normalised to the protein peak, it demonstrates that there is a relative and linear change in the relative water intensity compared to the normalised protein intensity with a change in water concentration i.e. with a decrease in phantom water concentration there is a decrease in the relative intensity of the water peak (Figure 2-10 B). In order to quantify this change in the ratio between the protein and water peaks, the water/total area ratio was calculated. A graph was then plotted to show the relationship between the water/total area ratio and known water concentration of the BSA/water solutions (Figure 2-11). It shows that the relationship is linear, however it is not directly proportional.

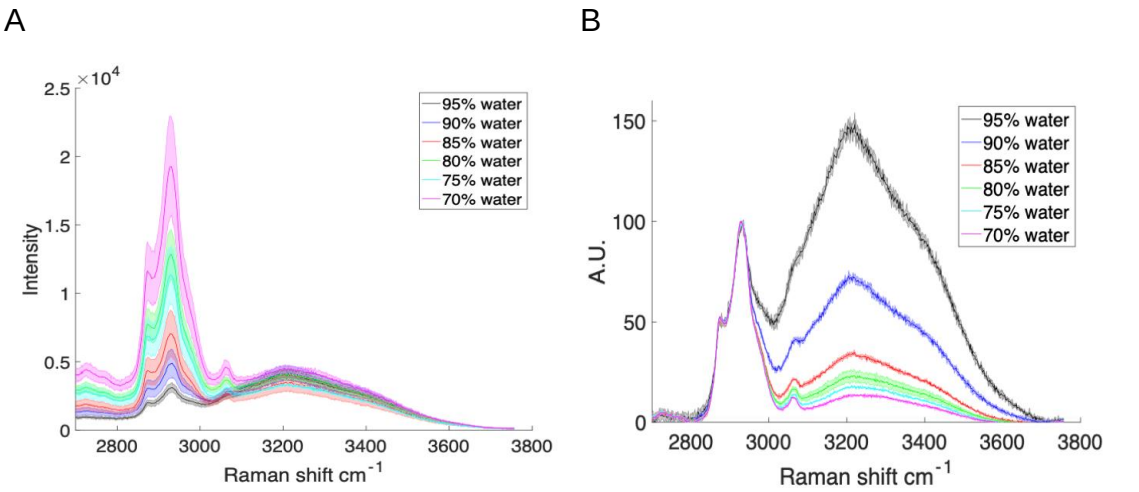


Figure 2-10 Graph showing Raman spectra in the HWN region of different concentrations of BSA/water solutions.

Data is mean of 3 repeats for each concentration and baselined using 1st order polynomial. A. raw spectrum – the peak intensity of the water peak between 3100-3700 cm^{-1} does not vary according to water concentration. B Spectral data normalised to protein peak at 2935 cm^{-1}- the relative intensity of water peak decreases corresponds with a decrease in water concentration. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial, shading in the same colour either side is +/- 1SD.

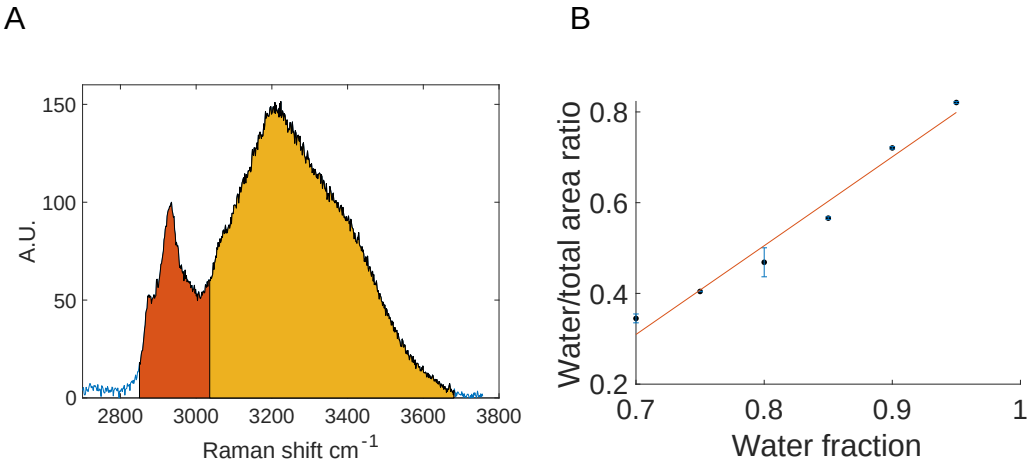


Figure 2-11 Graphs showing calculation of the water/total area ratio in BSA/water solutions with the InVia spectrometer.

A. Representative baselined Raman spectra of 95% water gelatine phantom showing the areas used for calculation of water concentration. Yellow area is the area under water curve of 3035 – 3680 cm^{-1}, orange area is the area under the rest of the HWN region 2850 – 3035 cm^{-1}. Area under water =yellow area; total area = yellow + orange area. B. Graph plotting mean water / total area ratio versus known water fraction of a number of different concentrations of BSA/ water solutions (n=3), Error bars +/- SD. A line of best fit showed a Gradient = 1.9576; RMSE = 0.0285

The suitability of the method of using the water/total area ratio to calculate the water concentration was tested. Gelatine phantoms of varying water concentrations were constructed, measured in a similar way, and the water/total area ratio calculated. The following equation was used to predict the water

concentration of the gelatine phantoms, using information on the relationship between water/total area ratio from the BSA/water solutions.

From Equation 2-2:

$$y = mx + c$$

Equation 2-2

This can be re-arranged to produce Equation 2-3 :

$$x = \left(\frac{y - c}{m}\right) X\ 100$$

Equation 2-3

Where x = predicted water concentration %, y = water/total area ratio of the gelatine phantom, m = gradient of the relationship between water/total area (1.9426), c = y intercept (-1.0609)

The results of this are in Table 2-5, and demonstrates that this method of calculation of water content is highly accurate when using substances of protein and water mixtures. It also demonstrates that the method for gelatine phantom construction produces phantoms of the desired concentrations.

WATER CONCENTRATION OF GELATINE PHANTOM (%)	WATER/TOTAL AREA RATIO	PREDICTED WATER CONCENTRATION %
95	0.79376838	94.8334218
90	0.71025914	90.5288215
85	0.61007738	85.3648132

Table 2-5 Table showing the validation of using the water/total area ratio to calculate water concentration.

The actual concentration of gelatine phantoms is based on the weight of water and gelatine, the calculated water/total area ratio is from the baselined spectra from Raman measurements. The predicted water concentration is calculated using the calibration graph from BSA data and Equation 2-3 to predict the water concentration of the gelatine phantoms.

Having established that the water/total area ratio having been calibrated to BSA/water solution was suitable for predicting water concentration in simple protein / water mixtures, the method was used in complex phantoms. Complex phantoms were constructed of intralipid, india ink, water and gelatine as described in 2.2.1.1. They were analysed and the data processed in a similar way to calculate the water/ total ratio. Equation 2-3 was used to predict the water concentration of the complex phantoms (Table 2-6), which shows that although there was a trend of decreasing predicted water concentration with decrease in

actual water concentration, it is inaccurate, and that BSA/water solution cannot be used to predict the water/total water area of complex phantoms.

NAME OF COMPLEX PHANTOM	ACTUAL WATER CONCENTRATION (%)	WATER/TOTAL AREA RATIO	PREDICTED WATER CONCENTRATION (%)
IL 50	78	0.504	79.9
IL 70	73	0.1796	63.1
IL 90	68	0.115	59.8

Table 2-6 Table showing the actual water concentration of intralipid phantoms, the calculated water/total area ratio, and the predicted water concentration based on the BSA/water calibration.

Reproducibility of breast phantom synthesis

If there is significant evaporation of water from phantoms during construction or measurement, it could lead to large variations between experiments and inaccuracies in measurements. If we are to regard the gelatine phantom as being a reference point for a particular water concentration, it needs to be reproducible and have a stable water content.

Three batches of gelatine phantoms of 95, 90 and 85% water concentration were constructed and measured at three different time points. Figure 2-12 shows there was no significant variation between experiments in the water/total area ratio suggesting the construction of gelatine phantoms is highly reproducible, and measurements within the same phantoms show minimal variation suggesting the phantoms are homogenous in nature.

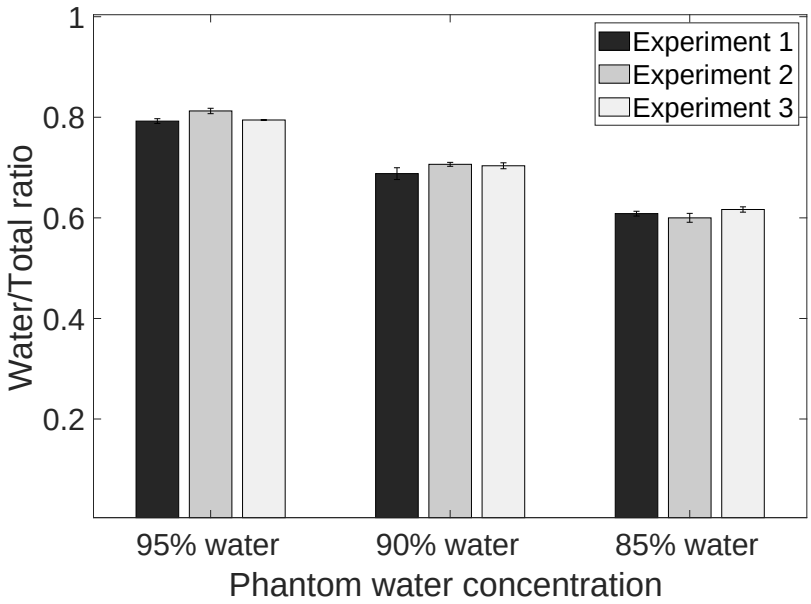


Figure 2-12 Graph showing gelatine phantoms of the same water concentrations in 3 different experiments and the water /total ratio.

Bars show mean water/total area ratio (n=3), error bars +/- SD. There was no significant difference in water/total area ratio between experiments (P>0.05;one- Way ANOVA).

During experiments, the gelatine phantoms can be on the lab bench for a number of hours. A gelatine phantom of 95% water concentration was measured at 0 and 4 hours and was at room temperature between measurements to demonstrate that significant changes in water concentration do not occur during this time (Figure 2-13).

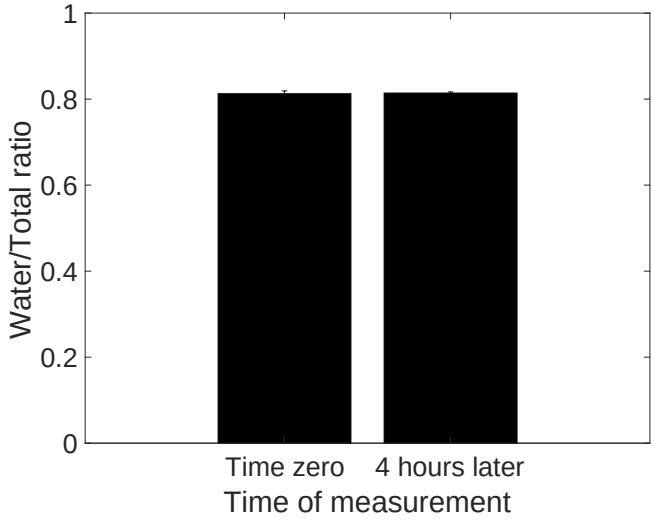


Figure 2-13 Graph showing the same gelatine phantom measured at time zero and 4 hours later.

Bars show mean water/total area ratio (n=3), error bars +/- SD. Student t-test demonstrated no significant difference in water/total area ratio between time points (P>0.05).

Discussion

These experiments demonstrate that the method for producing gelatine phantoms produces phantoms of the desired water concentration. It also demonstrates the suitability of the water/total area ratio for measuring changes in the water concentration as the relationship is linear in both protein/water and lipid/water mixtures. However, a protein/water mixture could not be used as a reference to predict water concentration in complex phantoms of lipid/protein/water. The reproducibility in construction and stability over time of gelatine phantoms is also demonstrated.

2.4 Evaluation of Raman systems

2.4.1 NP1- 680 nm laser excitation

2.4.1.1 Raman system set up

The general needle probe Raman system set up is described in 2.2.2. The following changeable components were used:

Laser excitation: 680 nm

Laser clean up filter: 680 nm

Camera: PIXIS 400 BRX

Measurement protocol: 1 second acquisition, 10 accumulations

2.4.1.2 Gelatine phantoms

The raw spectra obtained demonstrate a high background fluorescence at this wavelength (Figure 2-14 A). After pre-processing, it can be seen that with a higher water concentration there was a higher water peak compared to the protein peak (Figure 2-14 B). Figure 2-15 A shows the areas that were used to calculate the water/total area ratio. The water/total area ratio was calculated for all concentrations of gelatine phantom and plotted against known water concentration (Figure 2-15 B), which shows the relationship is linear, but not directly proportional.

A B

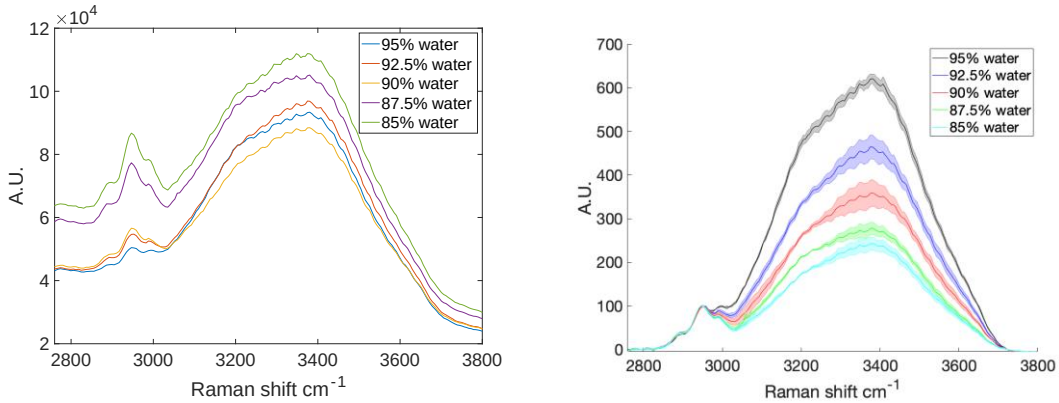


Figure 2-14 Raman spectra of a range of varying water concentrations of gelatine phantoms measured using NP1 Raman system.

A. Representative Raw Raman spectra at 5 different concentrations. B. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD.

A B

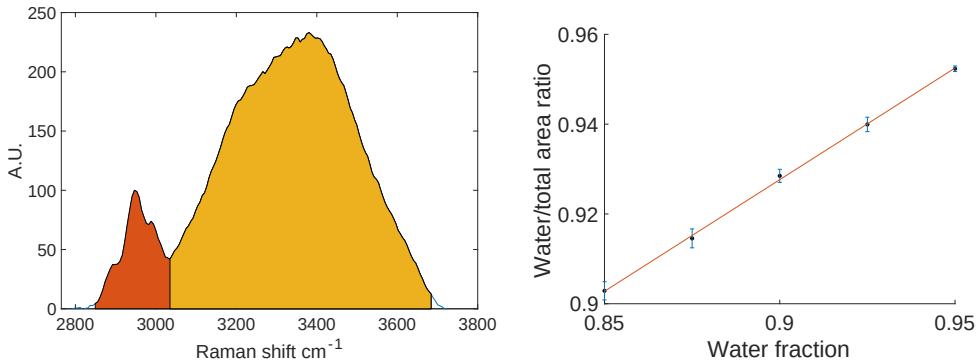


Figure 2-15 Using NP1 to measuring water/total area ratio in gelatine phantoms

A. Raman spectra of 85% water gelatine phantom with areas highlighted that are used to calculate water (yellow) / total (orange + yellow) ratio. B. Graph of water area/ total area ratio versus known water concentration of a number of different concentrations of gelatine phantoms measured using NP1 Raman system. Average ratio plotted (n=3), Error bars +/-SD. A line of best fit showed a Gradient = 0.4972; RMSE = 0.00048

2.4.1.3 Intralipid phantoms

Measuring intralipid phantoms in a similar manner, Figure 2-16 A shows that, in normalised spectra, with a decrease in known water content there is a decrease in water signal. It can be seen that the water peak has a different morphology to that of the water peak observed in the gelatine phantoms. This could be due to fluorescence in the intralipid phantoms. Calculation of the water/total area ratio to known water concentration demonstrates a linear relationship (Figure 2-16 B).

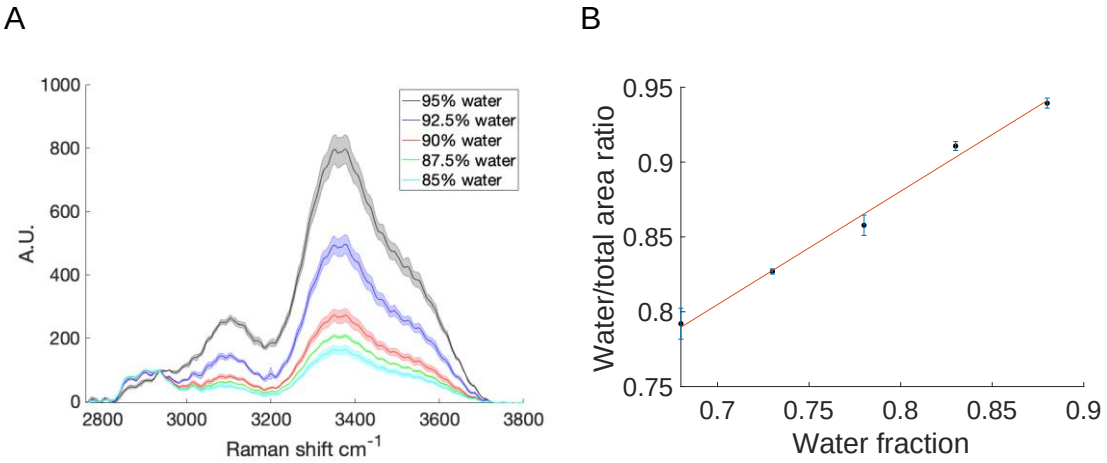


Figure 2-16 Demonstrating the ability to measure changes in water concentration in complex intralipid phantoms measured with NP1.

A. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD. B. Calculated water/ total area ratio versus known water concentration. Average ratio plotted (n=3), Error bars +/-SD, with a line of best fit. Gradient 0.7573, RMSE – 0.0050.

2.4.1.4 Recreating the challenges of the surgical environment

Pork meat and fat were analysed and compared to meat and fat that had been stained with blue dye or haemoglobin, burnt or wrapped in clingfilm. No Raman signal could be recovered from pork stained with blue dye or that had been burnt due to fluorescence that obliterated all signal. Haemoglobin staining and clingfilm covering did not make a visual difference to the Raman spectra (Figure 2-17), and comparison of water/total area ratios between pork meat vs Hb stained meat or clingfilm covered meat and between pork fat vs Hb stained fat or clingfilm covered fat revealed no significant difference (P>0.05; one- way ANOVA).

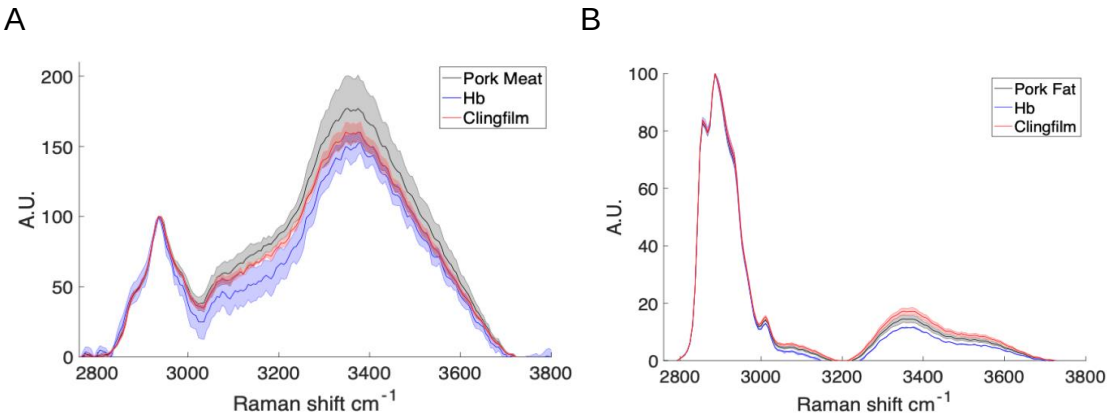


Figure 2-17 Results of measurements of pork phantoms with NP1.

Raman spectra of A. Pork meat and B. pork fat that had either not been treated, stained with haemoglobin or wrapped in clingfilm. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1 SD.

2.4.1.5 Discussion

These experiments with NP1 with 680 nm excitation demonstrate the ability of this system to obtain spectra in the HWN region. Using the water/total area ratio for both gelatine and complex phantoms shows a linear relationship, and demonstrates that NP1 can identify measurable changes in water concentration. Measuring pork tissue with haemoglobin and clingfilm covering did not significantly alter Raman spectra. However, blue dye and burnt tissue caused fluorescence that obliterated signal.

The large degree of fluorescence caused by burnt / carbonised tissue is a potential issue for analysing breast specimens. Diathermy artefact is usually a local process that would not prohibit analysis of an entire resected surface, but could ultimately reduce the sensitivity of NP1 analysis for IMA. The greater problem posed is the fluorescence caused by blue dye, which was significantly stronger than the Raman signal and resulted in no Raman scattering being detected. The intensity of fluorophores is related to the excitation wavelength, and it is generally highest in the UV or visible wavelengths [337]; by increasing the wavelength away from 680 nm, it is possible to move away from the fluorescence produced by blue dye. Using the laser excitation of 785 nm would minimise the influence of blue dye on Raman spectra, but would allow for measuring of the HWN spectrum within the confines of spectrometer limits (up to 1100 nm wavelength). Therefore the needle probe was re-configured for use with a 785 nm laser excitation to avoid the blue dye fluorescence.

2.4.2 NP-2 785 nm set up

2.4.2.1 Raman system set up

The general needle probe Raman system set up is described in 2.2.2. The following changeable components were used:

Laser excitation: 785 nm

Laser clean up filter: 785 nm

Camera: PIXIS 400 BRX

Measurement protocol: 1 second acquisition, 10 accumulations

2.4.2.2 Gelatine phantoms

The raw spectra obtained demonstrates the low intensity of Raman signal detected in the HWN region using this wavelength and a CCD camera (Figure 2-18 A). Despite this, after pre-processing, it was still possible to detect differences in water concentration; with a higher water concentration there was a higher water peak compared to the protein peak (Figure 2-18 B). A lower signal to noise ratio is noted. Figure 2-19 A shows the areas that were used to calculate the water/total area ratio. The water/total area ratio was calculated for all concentrations of gelatine phantom and plotted against known water concentration (Figure 2-19 B), which shows the relationship is linear, but not directly proportional.

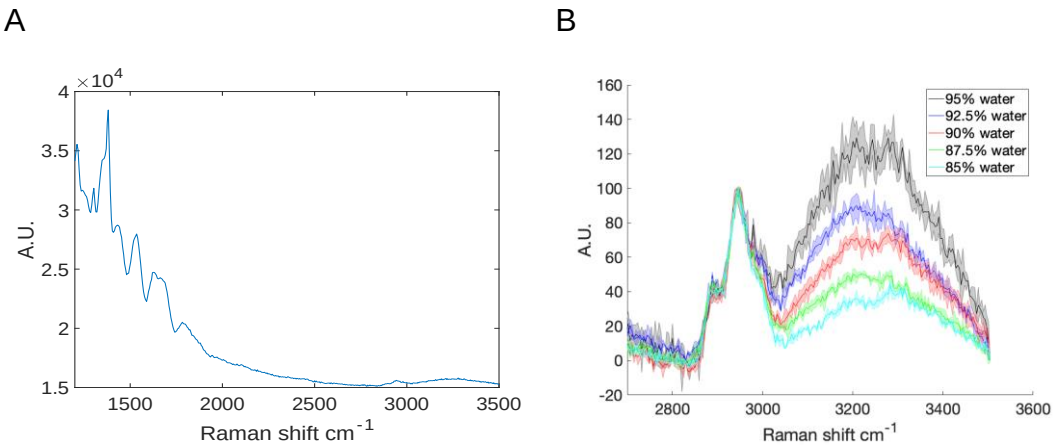


Figure 2-18 Raman spectra of a range of gelatine phantoms of varying water concentrations measured using NP2 Raman system.

A. Raw spectra from a 95% water gelatine phantom. B. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD.

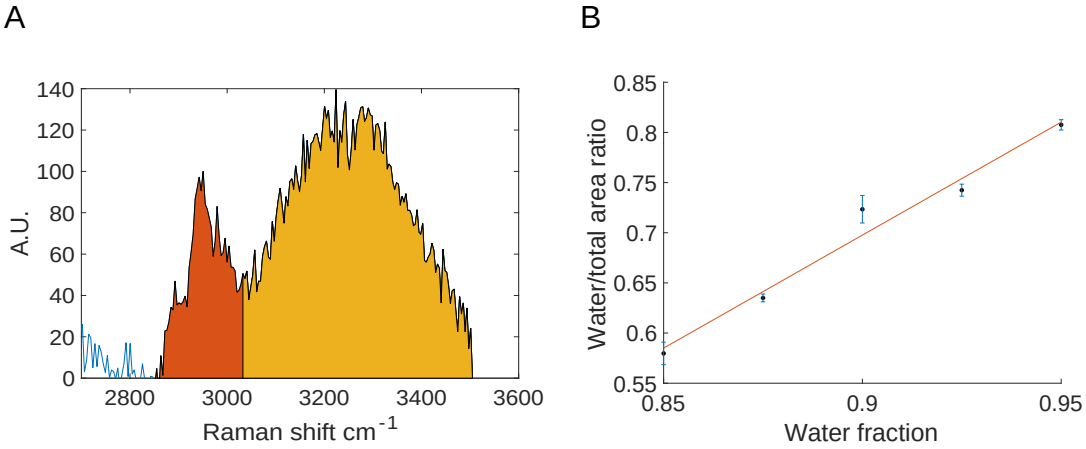


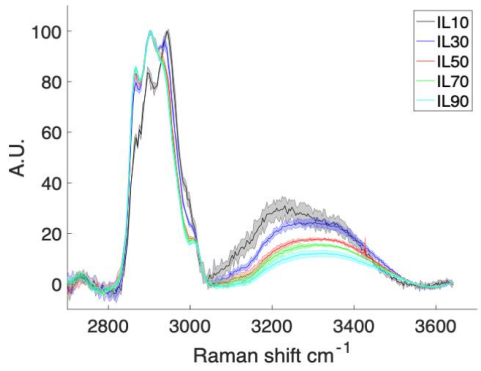
Figure 2-19 Using NP2 to measure water/total area ratio in gelatine phantoms

A. Raman spectra of 85% water gelatine phantom with areas highlighted that are used to calculate water (yellow) / total (orange + yellow) ratio. B. Graph of water area/ total area ratio versus known water concentration of a number of different concentrations of gelatine phantoms measured using NP2 Raman system. Mean ratio plotted (n=3), Error bars +/-SD. A line of best fit showed a Gradient = 2.2522; RMSE = 0.0132

2.4.2.3 Intralipid phantoms

Figure 2-20 A shows that, in normalised spectra, with a decrease in known water content there is a decrease in water signal. Calculation of the water/total area ratio to known water concentration demonstrates a linear relationship (Figure 2-20 B).

A



B

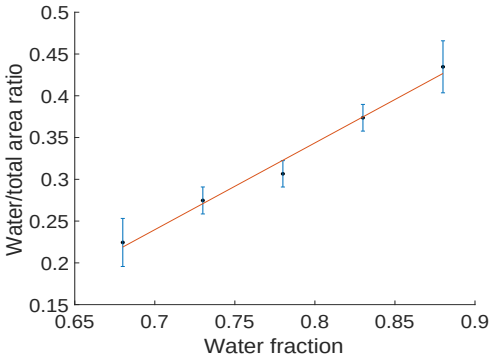


Figure 2-20 Demonstrating the ability to measure changes in water concentration in complex intralipid phantoms measured with NP2.

A. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD. B. Calculated water/ total area ratio versus known water concentration. Mean ratio plotted (n=3), Error bars +/-SD, with a line of best fit. Gradient 1.038, RMSE 0.0086.

2.4.2.4 Recreating the challenges of the surgical environment

Pork meat and fat were analysed and compared to meat and fat that had been stained with blue dye or haemoglobin, burnt or wrapped in clingfilm. Raman signal could not be recovered from pork that had been burnt due to fluorescence. After pre-processing the raw signal, blue dye staining, haemoglobin staining and clingfilm covering did not make a large visual difference to the Raman spectra (Figure 2-21), and comparison of water/total area ratios between pork meat vs blue dye stained meat or Hb stained meat or clingfilm covered meat and between pork fat vs blue dye stained meat or Hb stained fat or clingfilm covered fat revealed no significant difference (P>0.05; one-way ANOVA).

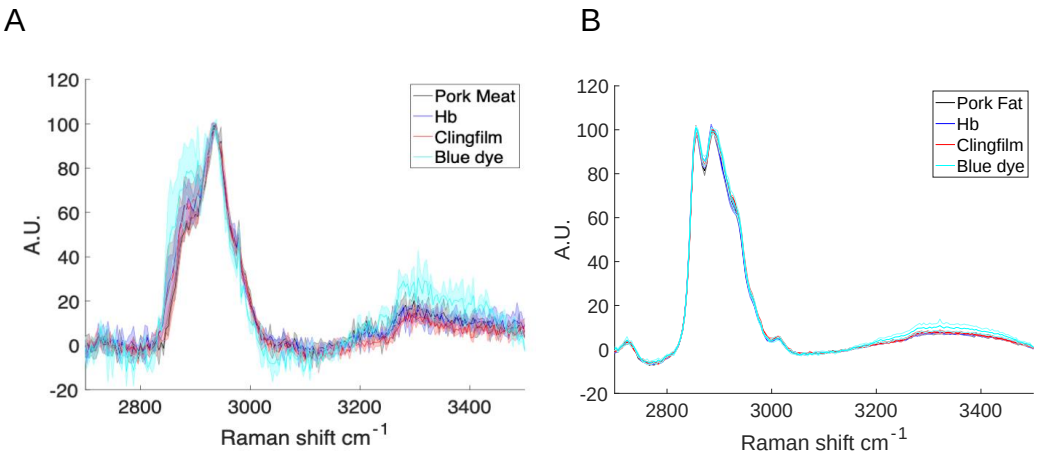


Figure 2-21 Results of measurements of pork phantoms with NP2 system.

Raman spectra of A. Pork meat and B. pork fat that had either not been treated, stained with blue dye, stained with haemoglobin or wrapped in clingfilm. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD.

Blue Dye Gelatine phantoms

To further investigate the effect of blue dye (BD) on obtaining Raman signal, blue dye gelatine phantoms containing concentrations of 0.01% and 0.1% blue dye were analysed using NP2 system. When comparing the raw spectra of gelatine phantoms with increasing concentrations of blue dye, it can be seen that there is an increasing background of fluorescence with blue dye. The Raman signal of gelatine can be recovered from this, but after pre-processing a 0.01% blue dye concentration alters the Raman signal, and led to a significantly different mean water/total area ratio (no BD 0.745 (SD 0.17), 0.01% BD 0.776 (SD 0.018), 0.1% BD 0.818 (SD 0.12); P=0.004; one-way ANOVA) (Figure 2-22).

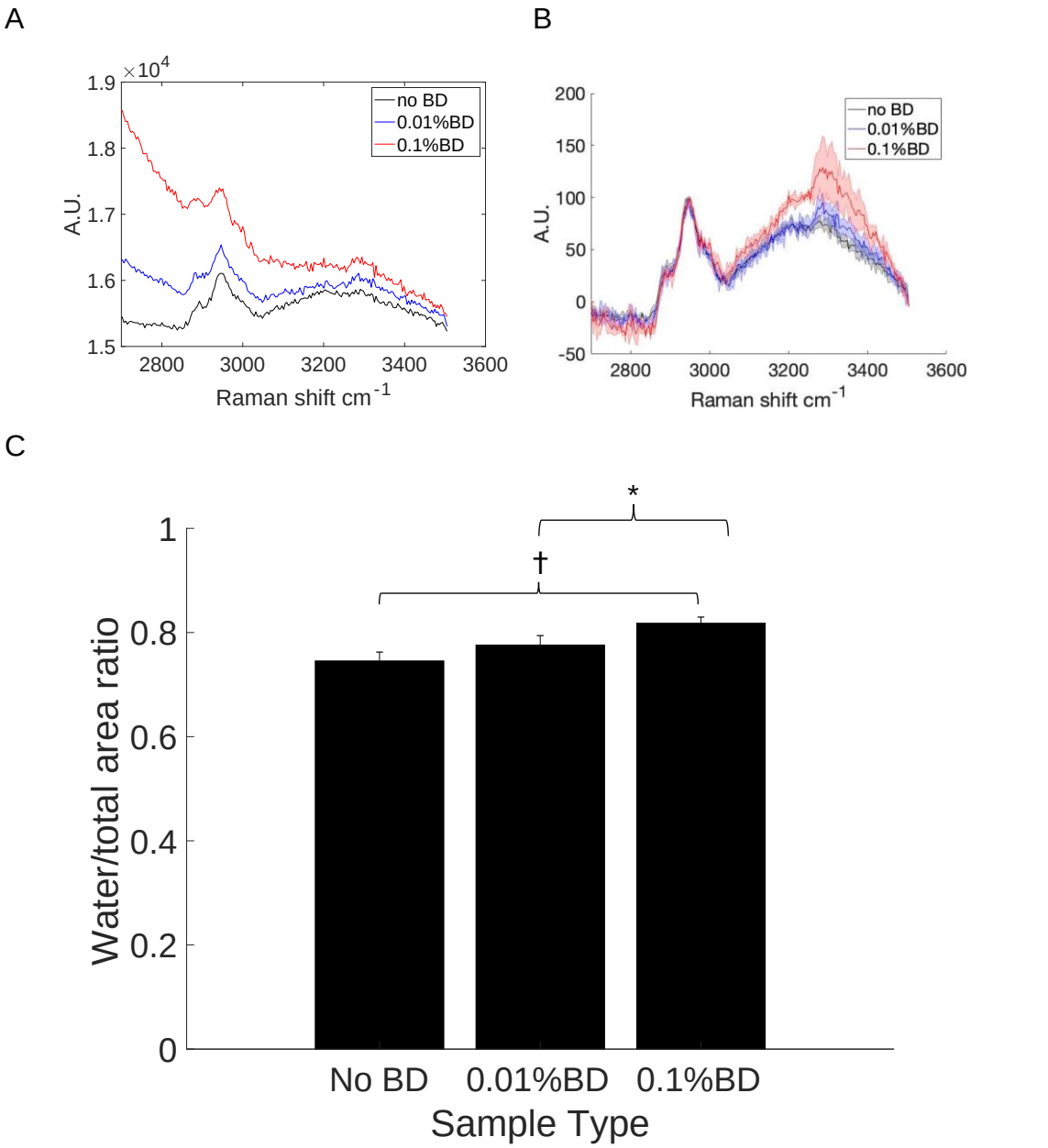


Figure 2-22 Results of measurement of blue dye (BD) gelatine phantoms with NP2.

A. Raw Raman spectrum with 10% gelatine phantom and 10% gelatine phantom with 0.01% and 0.1% blue dye. B. Raman spectra with after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD. C. Bar chart comparing the mean water/total area ratio between the three blue dye gelatine phantoms, *- significantly different ratio between 0.01% BD and 0.1% BD (P=0.043), † - significantly different ratio between no BD and 0.1% BD (P=0.0036); multiple comparison one-way ANOVA.

2.4.2.5 Discussion

The NP2 system is able to obtain Raman spectra in the HWN region. The system can also be used to quantify changes in water concentration in both gelatine and complex phantoms using the water/total area ratio as the relationship is linear. However when examining the raw spectra there are a number of issues identified

which are related to using a CCD camera to obtain HWN spectra using the 785 nm excitation wavelength.

There is low signal to noise and with a low signal intensity it could significantly affect the results as the high background fluctuations may alter the AUC, and subtle changes in water content at physiological levels could be immeasurable. The NP2 system is only capable of recording to a maximum Raman shift of around 3550 cm^{-1}, which is the cut off on all the graphs. And although it appears that the Raman spectrum tapers down at this extent – this is actually a result of the limitations of the CCD camera ability to read at this wavelength, rather than a true reflection of the Raman spectrum. If the NP1 spectra are compared with the NP2 spectra when measuring the same gelatine phantoms, it is obvious that in the middle of the water peak (at, for example, 3550 cm^{-1}) there is a high signal intensity with NP1, but zero intensity with NP2. This demonstrates that the Raman signals observed with NP2 are not a true reflection of the entire HWN region – particularly in the water peak region of 3035 – 3680 cm^{-1}, and that using the water/total area ratio, although a linear relationship is observed, the truncation of the water peak means it is not a true measurement of water concentration changes. The reason for these observed effects at the end of the CCD camera readings is due to the quantum efficiency of the CCD at this wavelength.

CCDs contain sectored pieces of silicon that allows accumulation of scattering over the whole exposure time, which allows discrimination between each frequency of scattered light and multiwavelength detection [100]. This allows for fast spectral scanning and rapid spectral acquisition; however silicon has an optimum region of sensitivity to photons (QE) and a detection cut off wavelength of 1100nm. For measuring in the HWN region and to investigate the water spectrum, the Raman system used must be able to detect Raman shift at 3035 – 3680 cm^{-1}. To understand this in relation to the quantum efficiency of the camera it has to be converted to the wavelength of light measured using Equation 2-4:

$$\lambda\,[nm] = \frac{1}{\left({}^{1}/_{\lambda_{ex[nm]}}\right) - \left({}^{Raman\ shift\ cm^{-1}}/_{10^{7}}\right)}$$

Equation 2-4 Conversion of Raman shift to wavelength Where λex = laser excitation wavelength; Raman shift = Raman shift of the molecule being investigated.

Using Equation 2-4, if using NP2, which is a 785 nm laser excitation to investigate a 3400 cm^{-1} Raman shift (the middle of the water peak):

$$\lambda\,[nm] = \frac{1}{\left({}^{1}/_{785}\right) - \left({}^{3400}/_{10^7}\right)}$$

$$\lambda = 1070\text{nm}$$

So the wavelength of 1070 nm is very close to the CCD cut off of 1100 nm. The manufacturer provides a quantum efficiency vs wavelength (nm) graph of the specific CCD camera used for NP1 and 2 (Figure 2-23). From this it can be seen that at a wavelength of 1070 nm, the quantum efficiency of the CCD is less than 10%, meaning the detection of Raman signal for the water signal is very poor, and suggests that NP2 is an unsuitable system for measuring changes in water concentration.

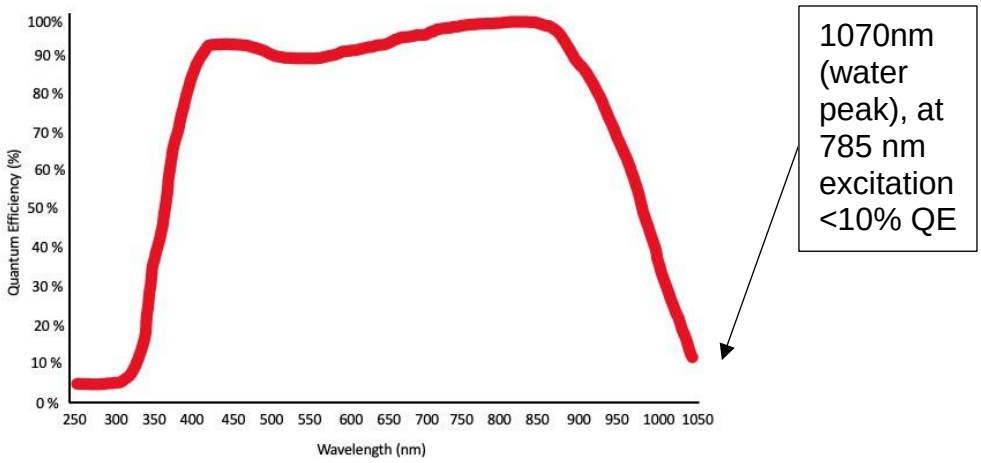


Figure 2-23 Quantum efficiency of the Princeton Instruments camera range versus wavelength. NP1 and NP2 used the PIX 400 BRX which is shown by the red line. Modified from reference [338].

The main reason for performing the experiments with NP2 was to assess if Raman spectra could be measured in the presence of blue dye. These experiments found that, unlike the NP1 data, Raman spectra could be measured in the presence of blue dye with NP2. Measurements from blue dye gelatine phantoms demonstrate that the raw Raman spectrum have an increased background fluorescence with blue dye, but Raman data is recoverable from this. This suggests that 785 nm excitation wavelength is appropriate for measuring the

HWN region in the presence of blue dye, if camera detection quantum efficiency can be improved. At a higher concentration of 0.1% blue dye in the gelatine phantoms, there was a significantly different water/total area ratio compared to a gelatine phantom without blue dye – however it may be that the low intensity Raman signal due to the poor camera QE in the HWN means the effect of fluorescence is more apparent, and it may be reduced with improved Raman scattering detection.

These findings suggest that a camera that uses a 785 nm excitation, but with a camera that has greater QE in the desired wavelength range may be the optimal system to measure the HWN region in the presence of blue dye. A solution to this is using a different type of camera to the silicon chip based CCD. Alternative detectors to silicon based CCD's are Indium gallium arsenide (InGaAs) detectors which are semiconductors with a lower bandgap to provide sensitivity in an extended near infrared range. The InGaAs camera available for use in this project is the Andor InDus InGaAs, the QE data for this camera is seen in Figure 2-24, which demonstrates that to investigate a wavelength of 1070 nm (middle of the water peak at 785 nm), the QE of this camera would be >85%. This is substantially better than the Princeton Instruments CCD camera, and so could be suitable for use with a 785 nm excitation wavelength for the investigation of water content in the presence of blue dye. This is tested in the next section – NP3.

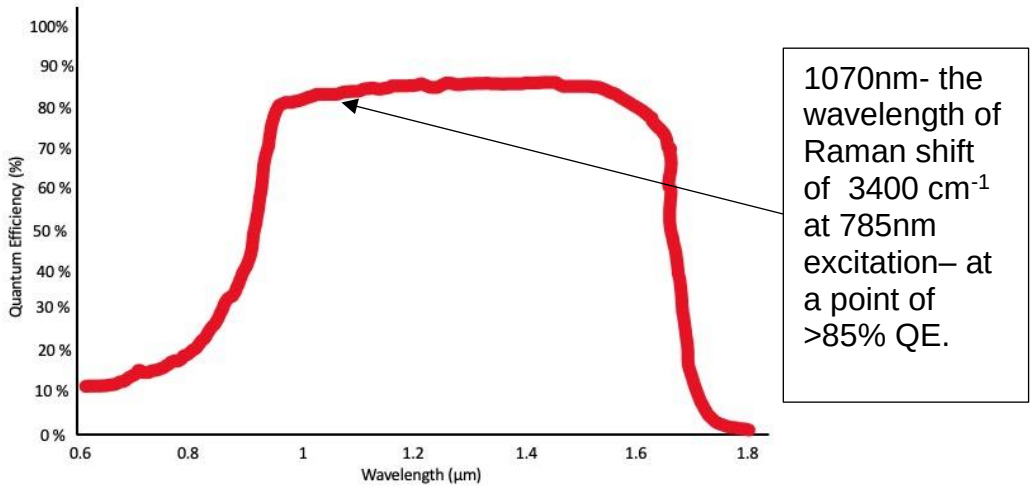


Figure 2-24 Quantam efficiency at 20°C of the Andor InGaAs camera range versus wavelength which is shown by the continuous red line. Modified from reference [339]

2.4.3 NP-3 785 nm InGaAs set up

2.4.3.1 Raman system set up

The general needle probe Raman system set up is described in 2.2.2. The following changeable components were used:

Laser excitation: 785 nm

Laser clean up filter: 785 nm

Camera: iDus InGaAs 1.7μm

Measurement protocol: 10 second acquisition, 5 accumulations

2.4.3.2 Gelatine phantoms

The raw spectrum of a 5% gelatine phantom, measured using NP-3 demonstrated in Figure 2-25 A shows that the signal is of much higher intensity than that gained with NP-2 and the CCD camera, and that the morphology of the water peak is similar to that of NP-1 and the water peak of the HWN region documented in the literature. A higher water peak intensity is seen with increasing water concentration when compared to the normalised protein peak (Figure 2-25 B). Figure 2-26 A shows the areas that were used to calculate the water/total area ratio. The water/total area ratio was calculated for all concentrations of gelatine phantom and plotted against known water concentration (Figure 2-26 B), which shows the relationship is linear, but not directly proportional.

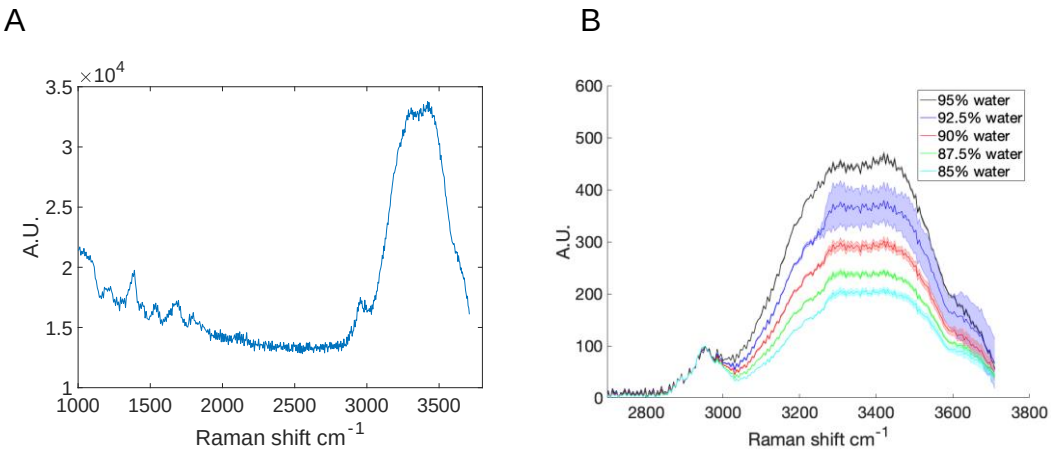


Figure 2-25 Raman spectra of a range of varying water concentrations of gelatine phantoms measured using NP1 Raman system.

A. Raw spectra from a 95% water gelatine phantom. B. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD.

A B

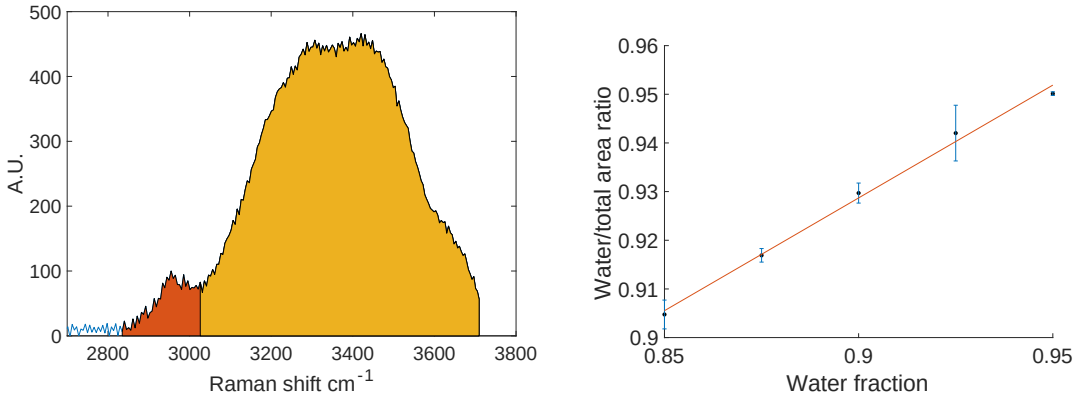


Figure 2-26 Using NP3 to measure water/total area ratio in gelatine phantoms

A. Raman spectra of 85% water gelatine phantom with areas highlighted that are used to calculate water (yellow) / total (orange + yellow) ratio. B. Graph of water area/ total area ratio versus known water concentration of a number of different concentrations of gelatine phantoms measured using NP3 Raman system. Mean ratio plotted (n=3), Error bars +/-SD. A line of best fit showed a Gradient = 0.46; RMSE = 0.0012

2.4.3.3 Intralipid phantoms

Figure 2-27 A shows that in intralipid phantoms with normalised spectra, with a decrease in known water content there is a decrease in water signal. Calculation of the water/total area ratio to known water concentration demonstrates a linear relationship (Figure 2-27 B).

A B

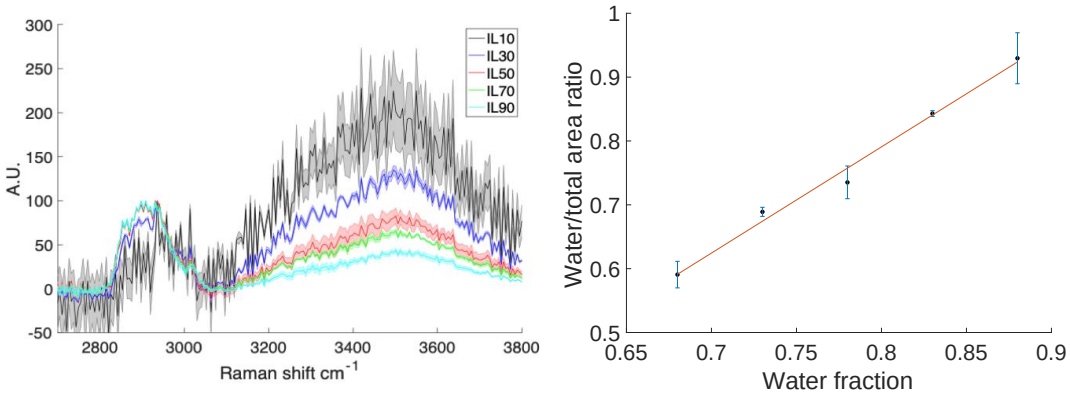


Figure 2-27 Demonstrating the ability to measure changes in water concentration in complex intralipid phantoms measured with NP3.

A. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD. B. Calculated water/ total area ratio versus known water concentration. Average ratio plotted (n=3), Error bars +/-SD, with a line of best fit. Gradient 1.75, RMSE 0.0164.

2.4.3.4 Recreating the challenges of the surgical environment

Pork meat and fat were analysed and compared to meat and fat that had been stained with blue dye or haemoglobin, burnt or wrapped in clingfilm. Similar to NP2, the Raman spectra were adversely affected by carbonisation in pork that had been burnt due to fluorescence, however the protein peak could be recovered, but the spectra from 3250-3700 cm^{-1} (water region) was altered and no Raman signal could be obtained in this region. After pre-processing the raw signal, blue dye staining, haemoglobin staining and clingfilm covering did not make a large visual difference to the Raman spectra (Figure 2-28), and comparison of water/total area ratios between pork meat vs blue dye stained meat or Hb stained meat or clingfilm covered meat and between pork fat vs blue dye stained meat or Hb stained fat or clingfilm covered fat revealed no significant difference (P>0.05; one-way ANOVA).

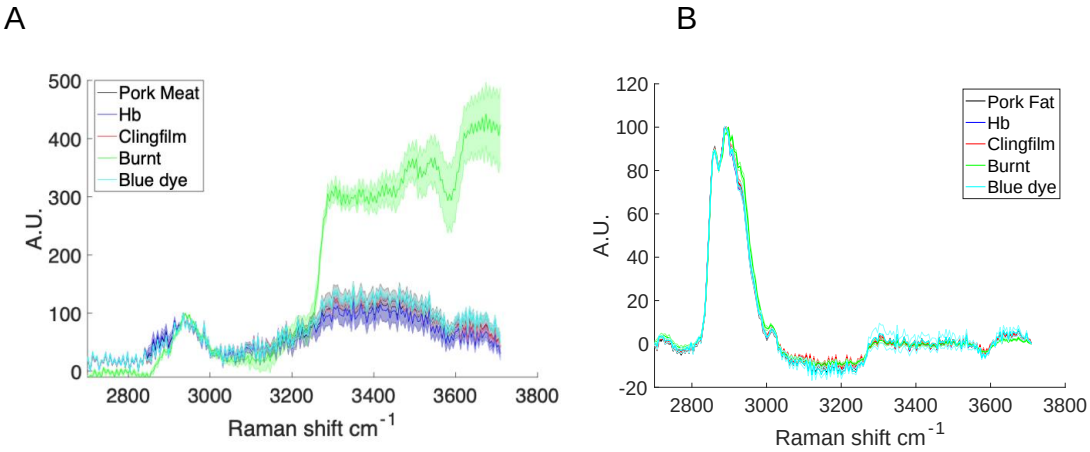


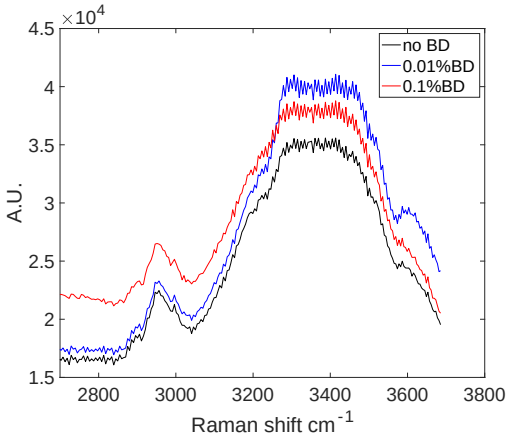
Figure 2-28 Results of measurements of pork phantoms with NP3 system.

Raman spectra of A. Pork meat and B. pork fat that had either not been treated, stained with blue dye, stained with haemoglobin, burnt or wrapped in clingfilm. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD.

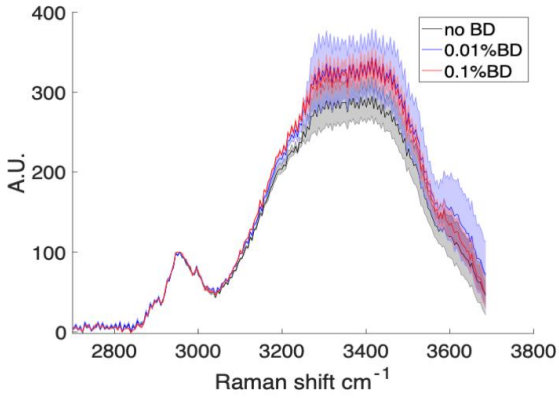
Blue Dye Gelatine phantoms

NP3 was used to analyse the blue dye gelatine phantoms. Similar to the spectra from the NP2 system, there was some increase in the fluorescent background in the presence of blue dye when measuring with NP3, however this appeared to affect the Raman spectra less than with NP2. There was no significantly different water/total area ratio (P>0.05; one-way ANOVA) between gelatine phantoms with 0%, 0.01% and 0.1% blue dye (Figure 2-29).

A



B



C

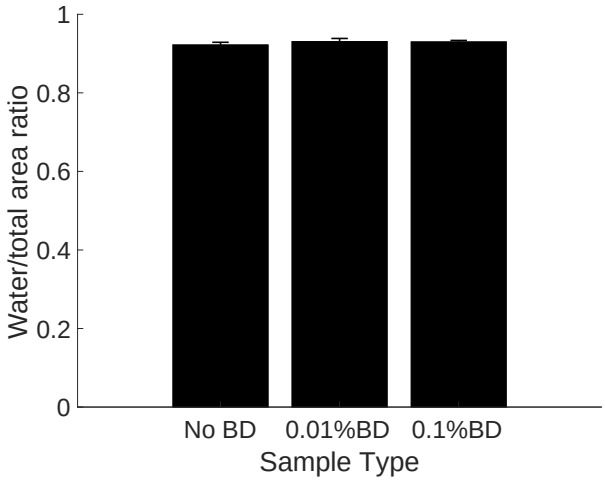


Figure 2-29 Results of measurement of blue dye gelatine phantoms with NP3.

A. Representative raw Raman spectrum with 10% gelatine phantom and 10% gelatine phantom with 0.01% and 0.1% blue dye. B. Raman spectra after pre-processing. Data is the mean (n=3) at each concentration after baselining using 1st order polynomial and normalised to the protein peak, shading in the same colour either side is +/- 1SD. C. Bar chart comparing the mean water/total area ratio between the three blue dye gelatine phantoms, statistical analysis showed no significant difference between mean ratios (P>0.05; one-way ANOVA).

2.4.3.5 Discussion

Measurements with the NP3 system confirm the ability to take high quality spectra in the HWN region, and using this system and the water/total area ratio, the water concentration can be measured. The water/total area ratios obtained for the different gelatine phantoms are very similar to those obtained with NP1, and the water/total area ratio to known concentration graph has a similar gradient

suggesting the information regarding water concentration is equivocal between NP1 and NP3.

The water/total area ratios from intralipid phantoms are markedly different between NP1 and NP3. The water spectra measured from NP3 from the intralipid phantoms has a morphology more consistent with what would be expected from the other experiments and from the literature [304, 333], compared to NP1. The background fluorescence of the intralipid phantoms was much greater in NP1, and it may be that the baseline subtraction was suboptimal, giving rise to the unusual morphology, which may have adversely affected the water/total area ratio values. Regardless of this discrepancy, the spectra gained from the intralipid phantoms from NP3 show a linear relationship between water/total area ratio and known water concentration.

The results from pork phantoms suggest that measurements from NP3 system are unaffected by blue dye, haemoglobin dye and clingfilm. Carbonisation does affect the signal and there is some fluorescence, however, a recognisable signal could be recovered which is improved from either the NP1 or NP2 system.
There was no difference in the water/total area ratio between gelatine phantoms with or without blue dye, this demonstrates the ability of NP3 system to measure Raman spectra accurately in the HWN region in the presence of blue dye with minimal disruption.

Despite these important improvements with NP3 compared to the other systems, the reduced signal to noise with the InGaAs camera is a potential issue. InGaAs cameras have high dark noise signal due to the low energy photons that are being detected[340]. This effect can be reduced by cooling of the detector (which was done to -85°C, as recommended by the manufacturer), but is still higher compared to silicon CCDs[339]. Due to this, longer acquisition times for spectral acquisition are required. In these experiments, overall measurement time was 5 times longer (50 seconds vs 10 seconds) with the InGaAs camera. This may have an impact in future work when taking spectra in a time-sensitive and clinical environment. With regards to margin analysis, it is an issue that will need optimisation for it to be effective. However, it is possible that these acquisition times could be reduced once the system spectral acquisition has been optimised,

and spectral analysis of breast tissue is more fully understood, to optimise the balance between spectral resolution and time for measurements to be struck for meaningful clinical data to be gained when measuring breast tissue. For initial, proof of concept experiments investigating HWN region and water concentration in breast tissue, NP3 has demonstrated its ability to measure water concentration and obtain spectra in the presence of blue dye, meaning it is the most suited system for this project.

2.5 Discussion

In this section of the thesis, the development of a Raman system capable of taking high quality HWN spectra, with appropriate analysis of the spectra to measure water concentration, whilst overcoming the potential barriers of taking measurements in the surgical environment has been explored. Tissue phantoms with a range of known water concentrations were constructed and analysed, and demonstrated to be reproducible in production and stable in water content over time. Using a closed laser, pre-calibrated, laboratory system to acquire Raman spectra from a standard of BSA solutions and applying this to the gelatine phantoms it was demonstrated that using a first order polynomial for baseline subtracting the data and measuring the water/total area ratio was an accurate method for determining water concentration.

A number of methods have been suggested for baseline subtraction in Raman spectroscopy and particularly in the HWN region. It is often performed using baseline subtraction using the visually most appropriate degree of polynomials, a number of different polynomials have been used including 1st [174, 271, 295], 2nd [337], 3rd [304] and 5th [341] order. More complex methods in the form of auto correction using multiple regression fitting have also been suggested [337]. The aim of baseline fluorescence subtraction is to reduce the influence of fluorescence on the data, without removing pertinent spectral information. In these experiments using phantoms, background fluorescence was not a major issue in the raw spectrum and baseline subtraction with a 1st degree polynomial visually appeared to be the most appropriate method. When using this method with the NP1 system which used a 680nm excitation, there was a higher degree of background fluorescence when measuring the intralipid phantoms, and it may be that the method of background fluorescence subtraction may have led to the unusual

appearance of the water spectrum seen. However, using other commonly used degrees of polynomials such as 2nd and 3rd order did not improve the results or visual appearance of the spectra, and it may be that more complex methods would be required in this particular case. However, as this system was not to be used in the project due to the major fluorescence from blue dye, this was not performed as it would not inform the project. It may be that measurements in breast tissue cause a higher degree of background fluorescence than the phantoms, and that different baseline subtraction methods are needed.

Spectral data from pure protein models (gelatine phantoms) regarding water/total area ratio could not predict water concentration in the more complex intralipid phantoms. The CH_3 stretch is responsible for both the Raman signals of protein and lipid, however the lipid band has a tendency towards higher levels of energy compared to protein[103]. The narrow protein peak of CH_3 is morphologically distinct to the broad lipid peak, and the intralipid phantoms were composed of both protein and lipid, and so a phantom that has the same water concentration, but a different protein/lipid ratio, will have a different area under the curve in the region 2850 – 3035 cm^{-1} which would affect the water/total area ratio. Although Masson *et al.* used this method in protein only gelatine phantoms, and extrapolated it to interpret water concentration in the protein/lipid rich cervix , our data suggests this is inappropriate [304]. To predict water concentration in fat rich breast tissue, water/total area ratio data from intralipid phantoms may be more appropriate.

All the systems were capable of measuring Raman scattering in the HWN region. NP1 and NP3 gave high quality, high intensity signals in this range, and NP2 gave lower intensity signals due to the quantum efficiency of the CCD camera. All three systems could measure the water concentration of the gelatine and intralipid phantoms. NP1 and NP3 gave very similar results in the measurement of gelatine phantoms and both produced HWN spectra of gelatine phantoms that were similar to the literature and both could detect changes in water concentration. NP2 spectral analysis produced quite different results, the water peak spectrum was truncated due to the quantum efficiency of the CCD camera in the HWN region, adversely affecting the ability of NP2 to measure water

concentration in the HWN region, meaning that NP2 was not suitable for this project.

The Raman systems were also evaluated for their suitability to obtain spectra in the presence of blue dye, haemoglobin, carbonisation and clingfilm, as this is essential for the future work of the project of taking measurements of surgically excised specimens. All three systems detected fluorescence in carbonised pork tissue; NP1 and NP2 could not recover any signal, and NP3 could measure the CH_3 region of protein, but the water signal was adversely affected. This is unsurprising as carbonisation intrinsically alters the breast tissue, and confounds analysis with the 'gold standard' of histopathology[319]. It may affect analysis in breast tissue, but it is likely to with any method of IMA, it is also usually localised to small areas rather than prohibiting analysis of an entire surface, and so an awareness of its effect on the Raman signal is required, but it does not preclude a Raman system from providing IMA.

Blue dye caused significant fluorescence in NP1, and no Raman scattering was detected in the presence in even a small amount of blue dye. Blue dye is used for sentinel lymph node detection in the majority of BCS, and staining of the breast tissue is common; therefore, a system that cannot obtain Raman spectra in the presence of blue dye is unsuitable for the preliminary work planned in this project, and for any clinically relevant method of IMA. Previous studies investigating HWN RS in breast tissue have not considered the effect of blue dye on the ability to obtain Raman spectra [299, 336]. These studies use laser excitation wavelengths (of 671 nm [299] and 532 nm [336]) that we predict would result in overwhelming fluorescence that obscures Raman signals in the presence of blue dye. Laser excitation with 785nm in NP2 and NP3 could obtain Raman spectra in the presence of blue dye. The Raman spectra, and subsequent analysis were more affected by blue dye in NP2, this may have been because the signal intensity was much weaker than in NP3, and so the slightly higher background fluorescence that the blue dye caused overpowered the weak Raman signal, whereas in NP3 there was a higher intensity Raman signal so the contribution of the fluorescent background was relatively much less.

2.6 Conclusion

NP3 which uses an InGaAs camera and 785 nm laser excitation has been shown to be the most appropriate system for assessing the HWN region in breast tissue for applicability in IMA, where blue dye is used. This is the first Raman system that has demonstrated the ability to obtain HWN spectra that are not significantly affected by the presence of blue dye, an essential attribute for a system to provide IMA. The NP3 system will be the Raman system used for the human tissue experiments in this thesis.

CHAPTER 3:
Investigating the ability of NP3 Raman system to quantify changes in water content using High Wavenumber Raman Spectroscopy

3.1 Introduction

In the previous chapter NP3 was demonstrated to be the most appropriate Raman system for investigating the provision of intraoperative margin analysis in human breast tissue. The analysis for this was performed using phantoms that allowed differentiation of performance regarding basic ability to evaluate changes in water content, and confounders that may occur in the clinical environment. The next stage of this system development is optimisation. This is both optimisation of the spectra that are obtained, and optimisation in the understanding of the data that can be obtained using the system. This chapter will demonstrate the method of optimal signal acquisition set up for NP3, evaluating the ability of NP3 to measure changes in water concentration at physiological concentrations and then describe a sequence of experiments on a range of phantoms to establish what changes are observed in the High Wavenumber (HWN) region spectra with a change in water concentration.

The intralipid breast phantoms measured in the experiments in Chapter 2 to ascertain the ideal Raman system for measuring HWN spectra in breast tissue had the limitation that the fat concentrations were not physiological. The intralipid phantoms had a range of fat concentration from 2-18%, with a water concentration being 68-88% whereas it has been suggested that the water concentration of breast tissue is in the range 30-60%, and correspondingly the fat concentration being 40-70% [281]. However, these figures are taken as global averages for an entire breast, or large areas of the breast. When measuring smaller samples of breast tissue, such as a wide local excision specimen for intraoperative margin analysis, one would see areas of almost exclusively fat,

and other areas exclusively of protein. Therefore, further investigation is performed in soya bean oil phantoms and in pork meat and fat to demonstrate the ability of NP3 to measure a change in water concentration at physiological ranges, which is necessary in order to be used in human breast tissue.

In section 2.2.4 it was noted that in a set of protein/water mixtures over a range of water concentrations, the spectral intensity of the maximum part of the OH peak (3390cm^{-1} in this case, due to the limitations of the CCD coupled with the InVia system used) did not change in relation to the change of water concentration. This spectral feature has been observed in other studies; using lyophilisation to produce cornea phantoms at a range of water concentrations, Bauer *et al.* [335] demonstrated that the OH peak Raman intensity (this time at 3400 cm^{-1}) did not change proportionally to the change in water concentration, and used a ratio of the maximum intensity of the protein peak to water peak to quantify the changes in water concentration. The conclusion drawn from this work has been cited as an example of how to use protein to water ratios to accurately determine water content, and the method has been validated since. However, the observation that the OH peak does not change in relation to water content, the possible significance of this and why this occurs is not addressed. If the water peak intensity doesn't change with changes in water concentration, and the water/total area ratio method is an accurate, validated method for quantifying changes in water concentration, it could be presumed that changes in the protein/lipid peak are proportionally changing with water concentration, though this has not been explicitly demonstrated previously.

In this chapter there is investigation and analysis of how these spectral changes may be interpreted to give a better understanding of the HWN region and will inform data analysis of measurements obtained from breast tissue in subsequent chapters.

3.2 Methods

3.2.1 Establishing the optimal signal acquisition time

The NP3 system uses an InGaAs camera, the range of Quantum Efficiency allows acquisition of spectra from the HWN region using a 785 nm excitation laser. An inherent limitation of InGaAs cameras is high dark noise compared to silicon CCDs [340] that are more commonly used for Raman Spectroscopy.
Noise has to be overcome in order to obtain a signal, and accounts for variation that can be seen between a number of identical measurements, which is expressed as the Signal to Noise Ratio (SNR) which is the average peak height above baseline divided by the standard deviation of the peak height [342].

Noise can come from a number of sources such as shot noise, fluorescence background noise, dark noise, and readout noise. Shot noise is the random variation in observed events, and so as the number of observed events increases (e.g. by increasing acquisition time), the shot noise decreases, and so the SNR increases. Dark noise is significant in InGaAs cameras. Dark noise is produced by spontaneous electron generation from the detector and is best demonstrated by obtaining a spectrum in a dark room without laser illumination[342]. In the presence of high dark noise lower acquisition times give a poor SNR, as the dark noise is proportionally more, and conversely by prolonging acquisition time more signal is obtained, with the same amount of dark noise, meaning the SNR improves with a longer acquisition time. Dark noise does not vary with laser intensity or variability in samples, and so there is the ability to subtract dark noise from spectra.

A method to improve the SNR in NP3 would be subtraction of fixed pattern noise as this is due to pixel to pixel variation within the InGaAs detector. Therefore, the detector dark noise was obtained by taking a measurement for the same acquisition time and number of accumulations as used for the sample measurement and a smoothed baseline obtained by Savitzky-Golay filter (7^{th} polynomial) as demonstrated in Figure 3-1.

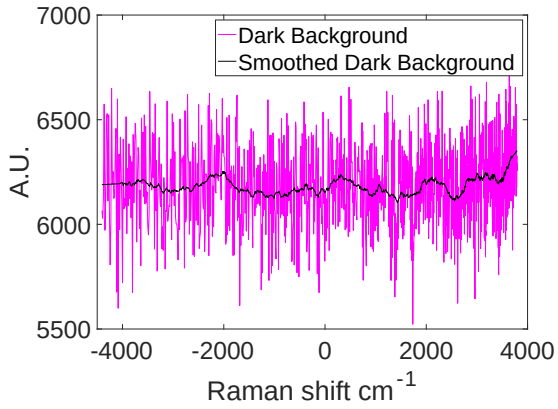


Figure 3-1 – Spectrum showing the Dark background and Smoothed Dark background at 1 second acquisition time with 5 accumulations.

The Smoothed Dark Background was obtained by Savitzky-Golay filter with a 7th order polynomial fit.

This smoothed baseline was then subtracted from the spectrum obtained from the sample, when measured for the same acquisition time and same number of accumulations as part of pre – processing. This gave improved spectral resolution and was successful at removing a significant source of noise from the spectrum, without affecting relevant spectral features (Figure 3-2).

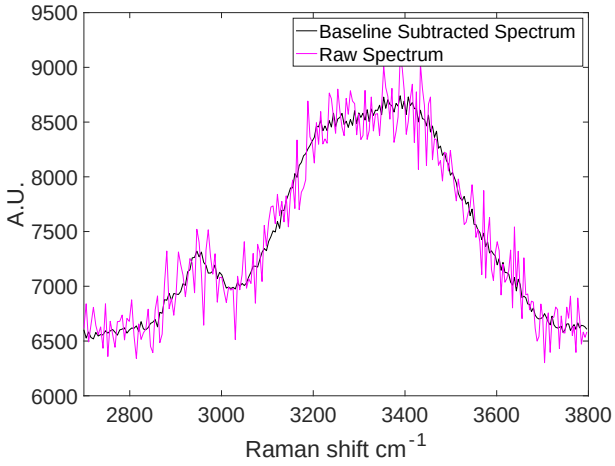


Figure 3-2 Figure showing the effect of Dark noise baseline subtraction on a HWN spectrum.

Representative spectrum of a Gelatine phantom of 90% water, 10% gelatine composition taken with 1 second acquisition time, 5 accumulations. Raw spectrum and spectrum after pre-processing with Dark noise 'baseline subtracted' shown.

Although longer acquisition times may improve the issue caused by shot noise, this may lead to sample heating degradation or burning, which for biological samples such as breast tissue, is to be avoided. A way to increase acquisition times and reduce the risk of burning is by taking shorter acquisition times with multiple accumulations, which may reduce sample heating if the laser turns off between acquisitions, allowing the sample to cool. However, the number of

accumulations taken may be limited by readout noise – that is the noise generated in the process of converting electrons detected into a digital signal to the computer. This noise increases with the number of accumulations.

A consideration for overall measurement time taken for measurement is that the time taken for the measurements needs to remain clinically relevant. If, in order to obtain satisfactory spectra, measurements took an hour each to obtain, neither the system nor the information gained from it would ever be useful in clinical practice. Therefore, we considered a maximum overall measurement time of 30 seconds per measurement to be one that was clinically relevant. Although this time per measurement is likely to be too long for use in practice, it is not so long as to be irrelevant. The area of collection of NP3 is 5.19×10^4 μm^2. With a view to NP3 providing IMA, for a lumpectomy sample of 60 X 60 mm (a normal size of sample) it would take approximately 720 individual measurements with NP3 to analyse a single resected edge. If each individual measurement took a second it would take 12 minutes to analyse a single resected edge, and if the individual measurement time took 25 seconds it would take 6 hours to analyse a single resected edge. So, although a measurement protocol of 30 seconds would not be clinically possible, the purpose of the experiments in this chapter are to explore subtle changes in spectra, which require higher spectral resolution. Once the HWN region of breast tissue is fully understood, it is likely the measurement time could be reduced to more clinically relevant times.

To find the optimal measurement settings of the number of acquisitions and accumulations required for NP3, measurements were taken with the NP3 system of a 90% water / gelatine phantom and performed pre-processing of baselining with a 1^{st} order polynomial and normalisation to the protein peak with dark noise baseline subtraction on the obtained spectra. Measurements were taken to compare the overall laser exposure time at 2 different times - 5 seconds and 25 seconds- to compare the signal to noise ratio of measurements. The exposure times were chosen to investigate if the spectra that were obtained in a short time period (5 seconds), were of sufficient quality to allow detailed analysis, as this time period would be closer to a clinically appropriate measurement time, or if a longer time period (25 seconds) was necessary.

Within these overall time periods, measurements were taken with different acquisition times and accumulations.

Results and Discussion

Figure 3-3 demonstrates the spectra obtained with an overall measurement time of 5 seconds, but with different acquisition and accumulation times. In both the 5 second X 1 accumulation and 1 second X 5 accumulations all peaks are clearly visible, however, there is more noise with 1 second X 5 accumulations, which may be due to an increased readout noise. Although this suggests that we could use an overall measurement time of 5 seconds, and this would still give reasonable quality spectra, we expect a significantly higher fluorescence background in biological tissue (as this has been observed in other studies investigating the HWN region in biological tissue [174, 304]), so longer acquisition times are likely to be necessary to obtain Raman features despite the higher background.

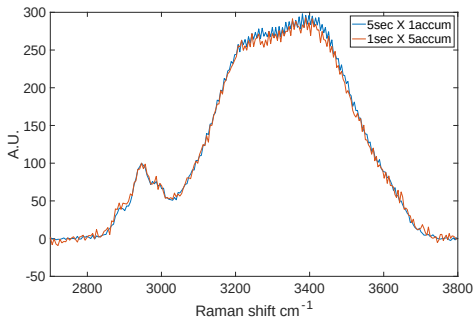


Figure 3-3 – Figure showing the difference in signal to noise in Raman spectra obtained with two different measurement protocols, both with an overall measurement time of 5 seconds.

Measurements obtained of a 90% water/10% gelatine phantom measured at 5 seconds X 1 accumulation (blue line) and at 1 second X 5 accumulations. Data is the sum of all accumulations to allow direct comparison and has been baselined, normalised to the protein peak and dark noise baseline subtracted.

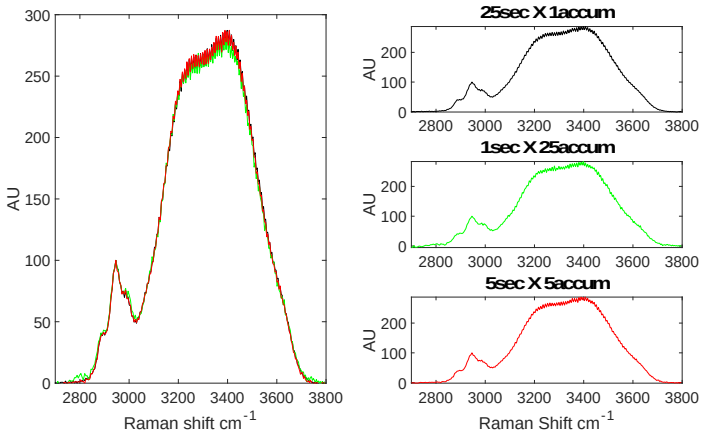


Figure 3-4- Figure showing the difference in signal to noise in Raman spectra obtained with three different measurement protocols, all with an overall measurement time of 25 seconds.
Measurements obtained of a 90% water/10% gelatine phantom measured at 25 seconds X1 accumulation (black line), at 1 second X 25 accumulations (green line), and 5 seconds X 5 accumulations (red line). Graph on left shows all three spectra combined, graphs on right show each spectrum separately to allow comparison. Data is the sum of all accumulations to allow direct comparison and has been baselined, normalised to the protein peak and dark noise baseline subtracted.

Figure 3-4 shows the results of spectra obtained at an overall measurement time of 25 seconds, with different acquisition times and accumulations. All spectra have improved spectral acquisition compared to the spectra obtained with an overall measurement time of 5 seconds. There are no obvious spectral differences between the three different measurement protocols, however, the spectra obtained with 1 second acquisitions with 25 accumulations has slightly increased noise which may be a result of a higher readout noise. Spectra obtained with a single 25 second acquisition produce spectra comparable to the spectra obtained with 5 seconds X 5 accumulations, however, multiple accumulations allows the laser light to be off between acquisition and may reduce biological sample heating and sample degradation.

Summary

Based on these results it can be considered that having a single acquisition time of 5 seconds rather than 1 second gives improved signal to noise, as would be expected. This is important as a smoothed spectrum may allow improved spectral analysis. Prolonged acquisition times could lead to sample degradation, so a measurement protocol of 5 second acquisition time with 5 accumulations with dark noise baseline subtraction is optimal based on the balance of signal to noise

and minimising the time taken to analyse a specimen, and so this measurement protocol is adopted in the measurements in the rest of this chapter.

3.2.2 Phantom Production and Measurement

3.2.2.1 Gelatine Phantom production

Gelatine phantoms were used to observe the changes in water concentration in the HWN region within a protein only environment. Gelatine phantoms were produced in the previously described method in Chapter 2 at concentrations of 85%, 87.5%, 90%, 92.5% and 95% water.

3.2.2.2 Soya bean oil/ water phantom production

Soya bean oil/water phantoms were used to observe the changes in water concentration in the HWN region with a lipid only environment and evaluate the ability of NP3 to measure the extremes of fat and water concentration that are likely to be encountered in clinical specimens. Olive oil would have been the preferred choice of oil, as it is composed of monounsaturated fatty acids and is predominantly oleic acid, which is the major fat in breast tissue [343, 344]. However, with 785 nm illumination, it has a broad fluorescent peak within the HWN region giving overwhelming fluorescence and no spectra could be obtained. Therefore, soya bean oil was chosen to create these phantoms, as it does not have this fluorescent band. Soya bean oil is predominantly polyunsaturated fat, composed of Linoleic Acid[344].

Soya bean oil phantoms were made according to a protocol adapted from a method by Merritt *et al.* [302] to make MRI suitable lipid phantoms. Organic soya bean oil (Clearspring,UK) was mixed with 4% of lipid volume of triton X100 (Sigma-Aldrich,Germany) which is a surfactant, which was added to allow emulsification of the lipid. This was warmed at 55°C for 5 minutes, before distilled water was added, mixed further, before sonification with a Hielscher Ultrasonics UP100H Handheld Ultrasonic Processor at 30kHz at 100% amplitude and 100% pulse (Hielscher Ultrasonics, Teltow, Germany) for 5 X 10 second pulses to create a liquid emulsion, and stored at 5°C until measurement at room temperature. A range of concentrations from 5% - 70% oil were made, (and 100% oil).

3.2.2.3 Pork serial dehydration

Serial dehydration with concurrent Raman spectra measurements was performed with pork to observe the changes in water concentration in the HWN region within a biological system, of both a protein rich (meat) and fat rich (fat) environment. Pork was purchased from a local supermarket (Sainsburys,UK), meat and fat was separated and cut into small cubes of <5 mm^3, and weighed. All specimens weighed less than 1 gram. Specimens were then weighed initially, and Raman spectra taken, and then left to air dry over the next 24 hours, within which at 7 different time points the specimens were weighed and Raman spectra taken. Raman measurements were taken at random points on the specimen's surface as it was not possible to mark points of previous measurement and take measurements from the same points at each measurement cycle. Changes in pork specimens were calculated as percentage weight change from the initial weight, and all changes in weight were assumed to be water evaporation and so provided an estimation of changes in water content of the specimen. There were negligible changes in the weight of the specimens after 24 hours (measured again at 36 and 48 hours), and so it was presumed that by 24 hours the maximum amount of water had evaporated using the air dry technique. The experiment was performed 3 times over 3 different days on 3 different pieces of pork, to account for possible variation in air drying/meat hydration. All experiments yielded similar results, data shown is from a single experiment to reduce the number of spectra/data points and allow better visualisation of results.

3.2.2.4 Mixed protein/lipid phantoms production

To investigate the effect of the protein/lipid environment on the HWN spectrum a set of gelatine based phantoms were created that had the same water concentrations (70 or 80%), but had varying protein/ lipid ratios. The conversion of oil weight to volume was based on the soya bean oil having a weight of 0.92 grams/ml. Triton X100 is a lipid, with a lipid spectrum in the HWN region the same as soya oil (and no protein or water peaks), and therefore the 'Total oil volume' was composed of the soya oil and Triton volume combined. This allowed the ratio of total protein and total lipid substrates that would contribute to the respective Raman spectra to be accurately calculated.

Name	Water (ml)	Gelatin (g)	Total oil (g)	Total oil (ml)	Soya oil (ml)	Triton (ul)
1	7	0.5	2.5	2.71	2.60	108
2	7	1	2	2.17	2.08	87
3	7	1.5	1.5	1.63	1.56	65
4	8	0.5	1.5	1.63	1.56	65
5	8	1	1	1.08	1.037	43
6	8	1.5	0.5	0.54	0.519	21

Table 7 - Table to show the composition of mixed phantoms composed of protein, lipid and water
In Phantoms 1 – 3 the water concentration was 70 %, and in Phantoms 4 – 6, the water concentration was 80%. Within these water concentrations, there were changing protein (gelatine) and lipid (soya oil) ratios.

3.2.2.5 HWN Raman experimental set up

HWN Raman spectra were acquired using the previously described NP3 set up (Chapter 2). Measurements were taken with a 5 second X 5 accumulation. 5 measurements were taken from different areas in each phantom, and acquired with Andor Solis (UK) software.

3.2.2.6 Data processing

Data was processed using Matlab. The mean was taken of the accumulations and then pre-processed by baselining using a 1st order polynomial, and dark noise background subtraction as described in section 3.2.1. The mean was taken of the 5 different area measurements, the standard deviation (if displayed) is of these 5 different area measurements. Normalised spectra were normalised to the highest peak in the spectra which related to the CH stretch of either protein or lipid (between 2850 -2950 cm^{-1}) [271, 333].

AUC water peak was calculated by taking the Area under the Curve of the Raman spectra between 3035-3680 cm^{-1}. And AUC CH stretch region (combined protein and lipid peak) was calculated by taking the Area under the Curve of the Raman spectra between 2850 – 3035 cm^{-1}. Water/total area ratio was calculated by taking the AUC water divided by AUC water + AUC CH stretch region[304].

Statistical comparison of water/total area ratio means was performed with one-way ANOVA, statistical significance set to $P<0.05$.

3.2.3 UV Vis experiments

UV-Vis absorbance of the individual components of the soya bean oil phantoms was investigated to assess how the optical absorbance characteristics of these components may affect the Raman spectra. If there is unequal absorbance by a constituent of the sample through the spectral range, a change in sample concentration may affect the Raman spectra due to absorbance as much as the effects of the change in sample concentration on the Raman spectra. This would affect the ability to quantify changes in the sample concentration, which was an area of investigation. UV-Vis absorbance spectra were obtained on a Thermo Scientific Evolution Array machine, using manufacturers software. Spectra were taken of soya bean oil, Triton X-100 and water with an integration time of 1 millisecond, average of 20 scans, with a 1 mm pathlength.

3.3 Results

3.3.1 Gelatine phantoms

The association between change in water content of gelatine phantoms and change in the water/total area ratio is shown in Figure 3-5 and was established in Chapter 2 – the relationship is linear and with a decrease in water content, there is a decrease in water/total area ratio. Measured with NP3, there is a gradient of 0.57.

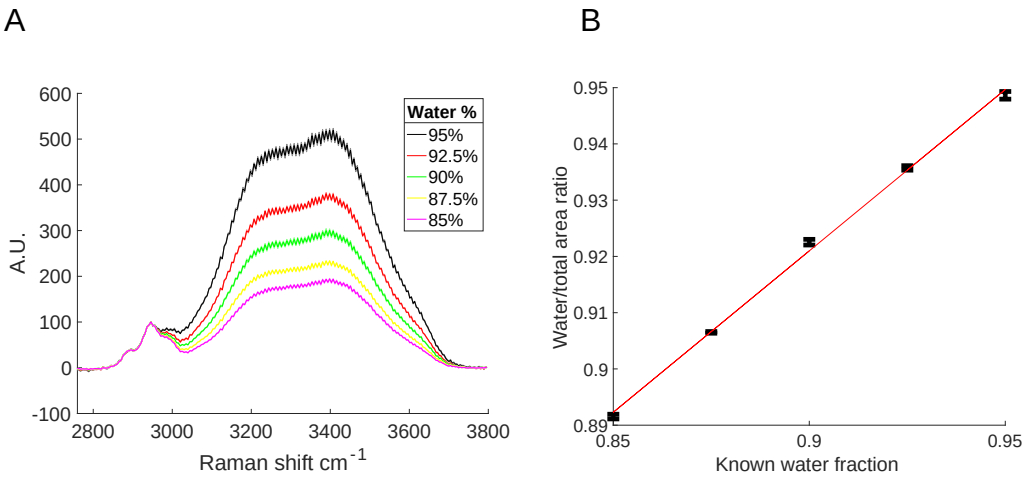
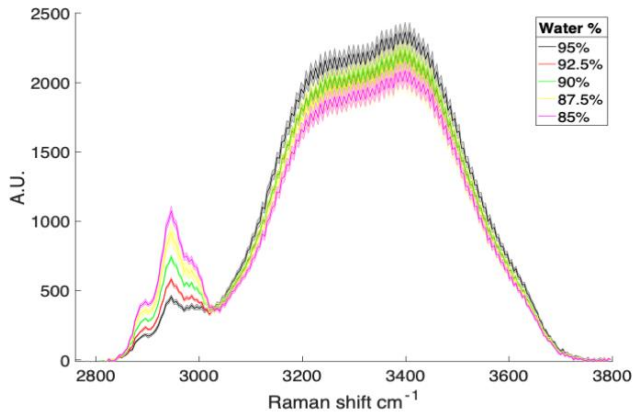


Figure 3-5 Figure showing Raman spectra in the HWN region of different water concentrations of gelatine phantoms and the relationship with water/total area ratio

A. Graph showing Raman spectra of gelatine phantoms at 5 different water concentrations, normalised to CH stretch region between 2850 -2950 cm^{-1}. Plotted lines are mean (n=5) for each concentration after pre-processing, shading in the same colour either side is +/- 1 SD. B. Scatter graph of known water fraction versus water/total area ratio. Points plotted are mean water/total area ratio (n=5), error bars +/- 1 SD. Red line is line of best fit (Gradient = 0.57,RMSE = 0.00092).

Figure 3-6 shows that in gelatine phantoms (a protein only environment), the water peak spectrum decreases with a decrease in water content, and the protein peak increases as the water content decreases. The Area Under the Curve (AUC) of the CH stretch region (corresponding to protein in these phantoms) increases proportionally with a decrease in water concentration (and increase in protein concentration) with low variability, and the Area Under the Curve (AUC) of the water peak decreases, however this is with a greater degree of variability and the error bars (representing standard deviation) overlap.

A



B

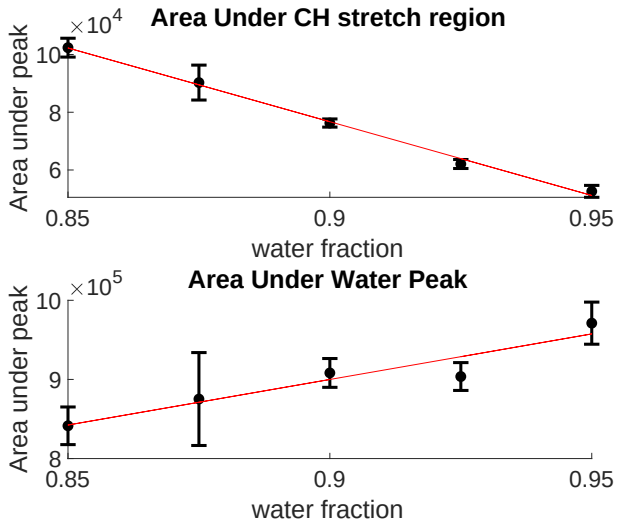


Figure 3-6 Graphs demonstrating the changes in the HWN region with changes in water concentration in a protein only environment.

A. Graph showing the spectral changes in protein peak and water peak signal intensity with changes in water concentration. Coloured lines are mean spectrum (n=5) after pre-processing, shading in the same colour either side is +/- 1 SD. B. Scatter plot demonstrating the change in Area Under Curve of CH stretch region (protein peak) and Area Under Curve of water peak at different water concentrations. Line of best fit to demonstrate linear relationship in change in water concentration and AUC of both CH stretch region (protein peak) and water peak.

All spectra had been pre-processed by baselining and dark noise background subtraction.

3.3.2 Soya bean oil phantoms

Figure 3-7 shows the normalised mean spectra of the soya bean oil phantoms ranging from 0% water to 95% water. It can be seen that, when normalised to the CH stretch region (corresponding to a lipid peak in these phantoms) at 2940cm^{-1}, the water peak decreases with a decrease in the water concentration of the phantom. However, the difference seems greater in the phantoms with high water concentrations. The scatter graph demonstrates that there is a linear

relationship between known water concentration of the phantoms and the water/total area ratio. If the phantoms are grouped and analysed according to those with higher water concentration (70-95% water) and lower water concentration (30-65% water), there is a slight difference in the gradients (higher water concentration gradient= 1.61 vs lower water concentration gradient = 1.04). This suggests that at a lower water concentration (and a higher lipid concentration) the difference in water/total area ratio between different water concentrations reduces, and so the ability to differentiate between different water concentrations may be reduced.

A B

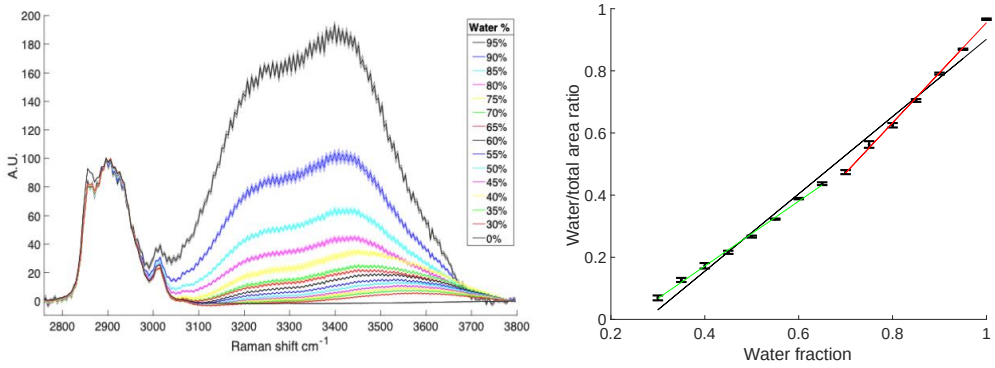


Figure 3-7. Figure showing Raman spectra in the HWN region of different water concentrations of soya bean oil phantoms and the relationship with water/total area ratio

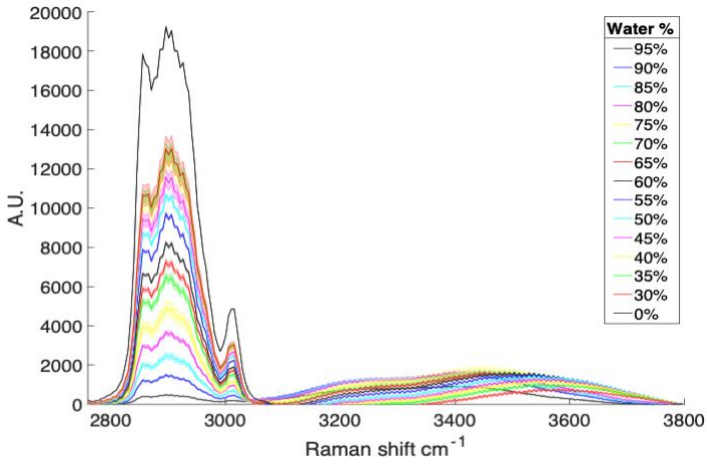
A. Graph showing Raman spectra of soya bean oil phantoms at 15 different water concentrations, normalised to CH stretch region between 2850 -2950 cm^{-1}. Plotted lines are mean (n=5) for each concentration after pre-processing, shading in the same colour either side is +/- 1 SD.

B. Scatter graph of known water fraction versus water/total area ratio. Points plotted are mean water/total area ratio (n=5), error bars +/- 1 SD. Black line is line of best fit of all data points (n=15) (Gradient=1.24, RMSE = 0.031), red line is line of best fit of high water concentration phantoms (70-95% water) (Gradient = 1.61, RMSE = $4X10^{-16}$), green line is line of best fit of low water concentration phantoms (30 – 65% water)(Gradient = 1.04, RMSE=$1.3X10^{-16}$).

Figure 3-8 shows that in soya bean oil phantoms (a lipid and water only environment), the water peak spectrum does not appear to decrease with water content, and the lipid peak increases as the water content decreases. The scatter graph shows that with a decrease in water concentration (and subsequent increase in lipid concentration), the AUC of the CH stretch region (lipid peak) increases with a linear relationship, with narrow error bars suggesting a small degree of variability in measurements. However, the AUC of the water peak

shows no consistent relationship with a change in water concentration. At higher water concentrations (water fraction 0.60-0.95) the AUC of water peak changes do not have any pattern with a change in water concentration, whereas at lower water concentrations (water fraction 0.55-0.3) with a decrease in water concentration there is a decrease in AUC of the water peak.

A



B

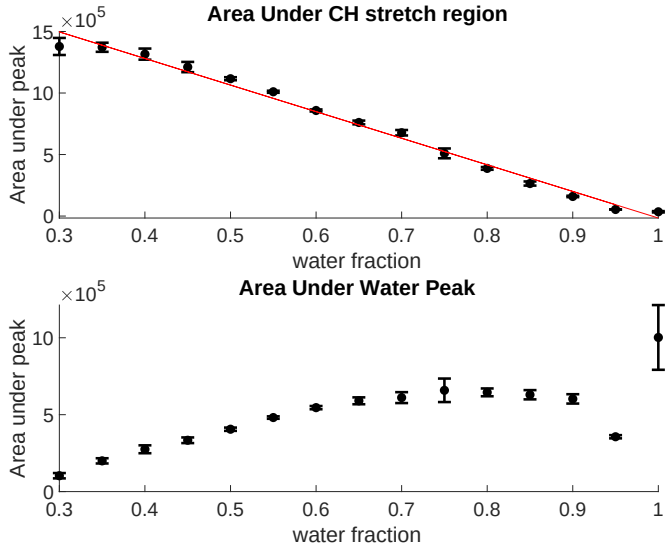


Figure 3-8 – Graphs demonstrating changes in the HWN region with changes in water concentration in a lipid only environment.

A. Graph showing the spectral changes in lipid peak and water peak signal intensity with changes in water concentration with a range from 30 – 95% water. Coloured lines are mean spectrum (n=5) after pre-processing, shading in the same colour either side is +/- 1 SD. B. Scatter plot demonstrating the change in Area Under Curve of CH stretch region (lipid peak) and Area Under Curve of water peak at different water concentrations. Line of best fit to demonstrate linear relationship in change in water concentration and AUC of CH stretch region (lipid peak) only.

All spectra had been pre-processed by baselining and dark noise background subtraction.

3.3.2.1 UV-Vis results

To assess the influence of optical properties on the HWN region spectra, absorption spectra were taken from the soya bean oil phantoms. These were the only phantoms that were suitable for UV-Vis analysis.

Figure 3-9 shows there is little change in the absorption of soya bean oil and Triton X100 through the relevant range of wavelength. However, there is some higher absorption of water at the lower wavenumbers, which could disproportionally affect the spectrum and the relevant peak ratios, particularly if the path length of the Raman scattered photons is increased or decreased by the presence of the highly scattering lipid droplets to pass through more of the phantom.

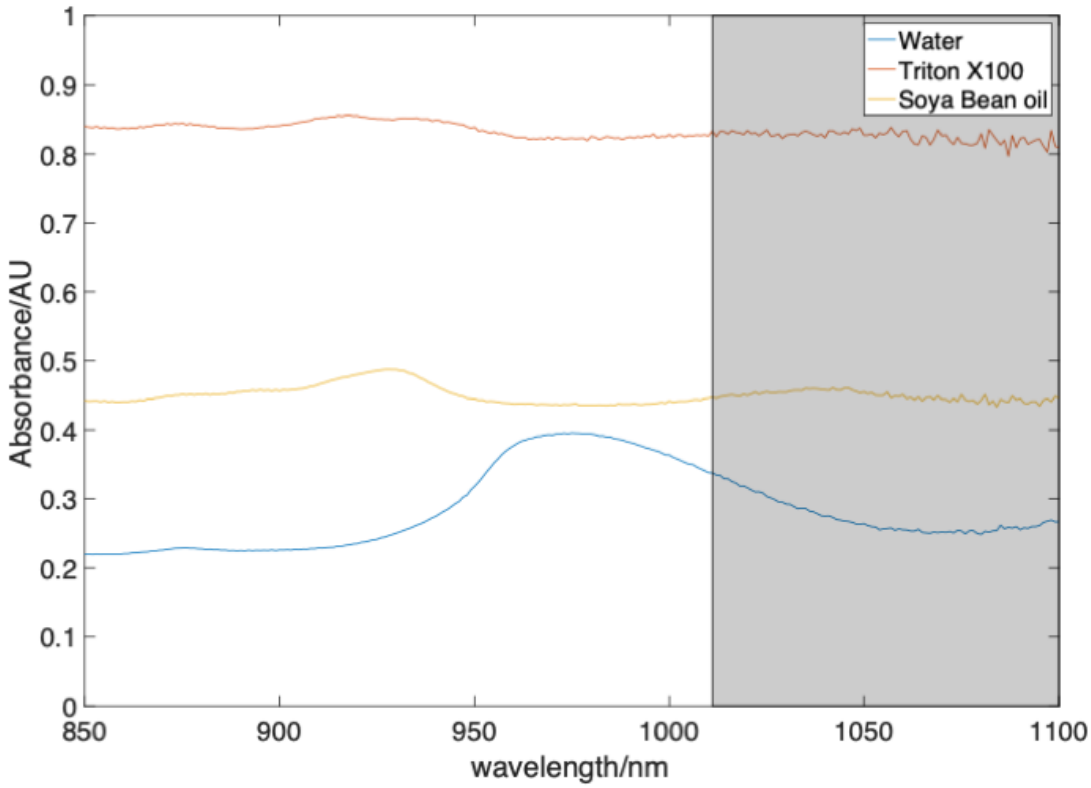


Figure 3-9- Graph demonstrating the UV Vis absorption spectra of water, Triton X100 and soya bean oil

Measurements were taken in a cuvette with 1mm pathlength for 1 millisecond with mean of 20 accumulations. Shaded area demonstrates the wavelength investigated in the HWN region equivalent to Raman shift 2850 – 3650 cm^{-1}.

3.3.3 Pork phantoms

3.3.3.1 Meat

The serial dehydration of pork specimens over 24 hours led to 60% change in the meat specimen weight. Figure 3-10 shows the spectral results from the serially dehydrated pork meat- as pork meat reduced in weight (which is presumed to be predominantly dehydration and decreasing water concentration), there was a reduction in the water/total area ratio. The scatter graph shows that this relationship is not linear, as the pork changed from 100% weight to 77% weight there was very little change in the measured water/total area ratio, and after this the water/total area ratio reduced with weight reduction.

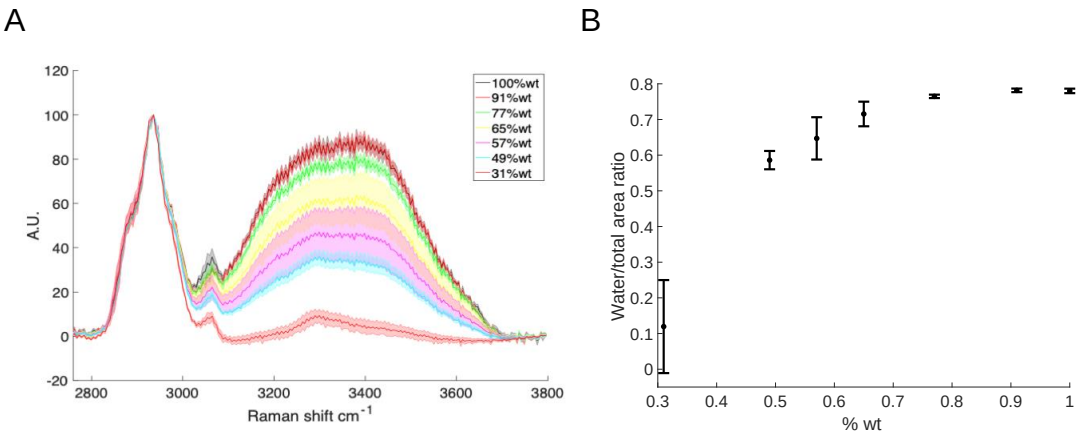


Figure 3-10 Figure showing Raman spectra in the HWN region of pork meat at different stages of dehydration and the relationship with water/total area ratio

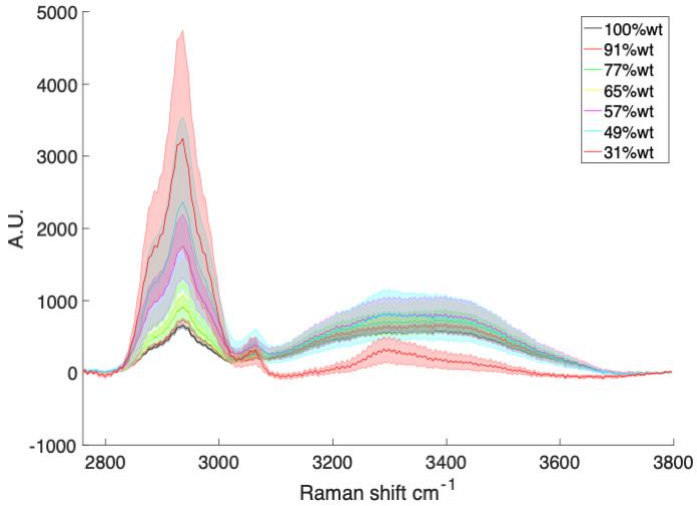
A. Graph showing Raman spectra of pork meat specimens at 7 different stages of dehydration (and different water concentrations). Plotted lines are mean (n=5) for each %age change in weight normalised to CH stretch region between 2850 -2950 cm^{-1}, after pre-processing, shading in the same colour either side is +/- 1 SD.

B. Scatter graph of %age change in weight versus water peak/total area ratio. Points plotted are mean water/total area ratio (n=5) from spectra taken at each stage of dehydration, error bars +/- 1 SD.

Figure 3-11 demonstrates that in a biological system of pork meat (a protein rich environment), the water peak spectrum does not appear to decrease with water content, and the CH stretch region (relating to protein in this protein rich environment) increases as the water content decreases. The scatter graph shows that with a decrease in water concentration the AUC of the CH stretch region increases with a linear relationship. However, the amount of protein is unchanged. It also shows that the AUC of the water peak does not have a consistent relationship with a change in water concentration. There is a trend that the AUC of the water peak increases with a decrease in water concentration,

though the error bars are wide. This is true until the final measurement, where the specimen is very dehydrated, when the AUC water peak decreases.

A



B

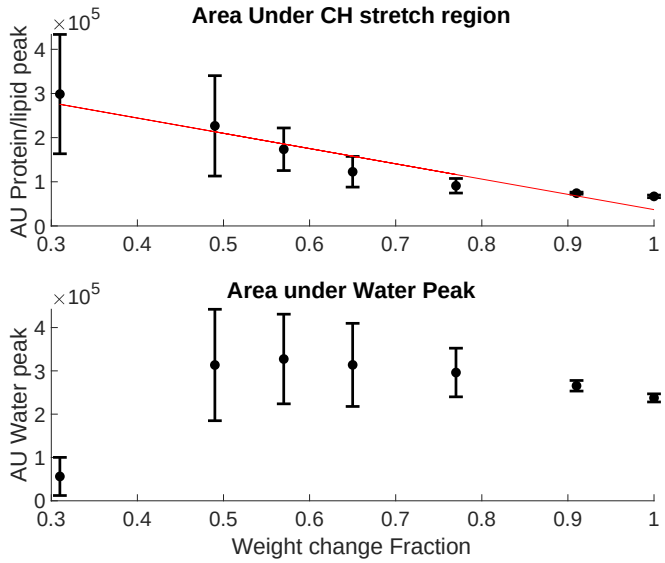


Figure 3-11 – Graphs demonstrating changes in the HWN region with changes in water concentration in a biological system.

A. Graph showing the spectral changes in CH stretch region (protein peak) and water peak signal intensity with changes in water concentration with a range from 100% weight – 45% weight. Coloured lines are mean spectrum (n=5) after pre-processing, shading in the same colour either side is +/- 1 SD. B. Scatter plot demonstrating the change in Area Under Curve of CH stretch region (protein peak) and Area Under Curve of water peak at different water concentrations. Line of best fit to demonstrate linear relationship in change in water concentration and AUC of CH stretch region only.

All data has been pre-processed by baselining and dark noise background subtraction.

Water in biological systems are sequestered in different environments, broadly water is bound or unbound (free). It is known that the water peak in the HWN region is composed of peak contributions from water sequestered in different environments [336]. To analyse the exact contributions of each different water environment requires detailed analysis using Voigtian curve fitting [304], however a basic assessment of changes in the water environments can be performed by looking at the spectral peaks intensities of the different environments [334]. Using this method, the water peak between 3035 – 3680 cm^{-1} can be analysed by comparing the spectral intensities of contributions from bound water (corresponding to the peak at 3300 cm^{-1}) and unbound water (corresponding to the peak at 3420 cm^{-1})[304, 334, 345].

To investigate changes in the types of water within the serially dehydrated pork meat, the change in ratio between bound water versus unbound water signal intensity was calculated. Figure 3-12 shows the spectral features that were used to calculate the ratio and demonstrates that with greater dehydration (and a lower water concentration), the proportion of unbound water decreases.

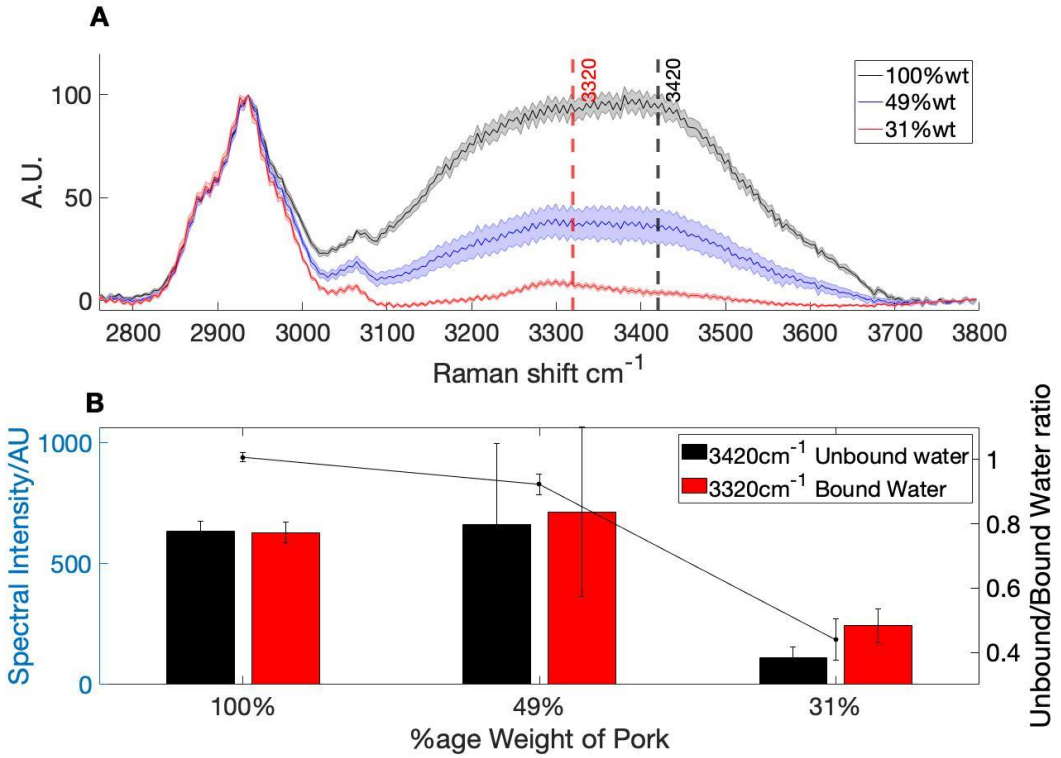


Figure 3-12. Graphs showing the change in Unbound versus Bound water in serially dehydrated pork.

A.Shows HWN Raman spectra of serially dehydrated pork at three different points of dehydration, with a range from 100% weight – 45% weight. Plotted lines are mean (n=5) for each %age change in weight normalised to CH stretch region between 2850 -2950 cm^{-1}, after pre-processing, shading in the same colour either side is +/- 1 SD. Dotted lines show the spectral features of Unbound water at 3420 cm^{-1} (black), and the bound water at 3320 cm^{-1} (red). B. Shows the Spectral intensity on left axis of Unbound and bound water demonstrated with bar graphs, with unbound/bound water ratio on right axis demonstrated with line graph at different %age weight pork. All values plotted are mean (n=5), error bars +/- SD

3.3.3.2 Pork fat

The serial dehydration of pork fat over 24 hours led to an 18% change in the fat specimen weight, compared to the 60% change in the meat specimen weight. This suggests that pork fat has a reduced amount of initial water than pork meat. The results from the serial dehydration of pork fat are shown in Figure 3-13. This shows that there was very little change in the water/total area ratio with weight change. The relationship between percentage weight change and water/area total ratio demonstrated in graph B is along the y axis of 0, the reason for negative values was that baselining in pork fat spectra was difficult as there was an inflection at the end of the spectrum due to tissue fluorescence, that was not possible to account for leading to the occasional negative reading within the water

spectrum (as the baseline fluctuated around zero), which then gave a small negative water area figure.

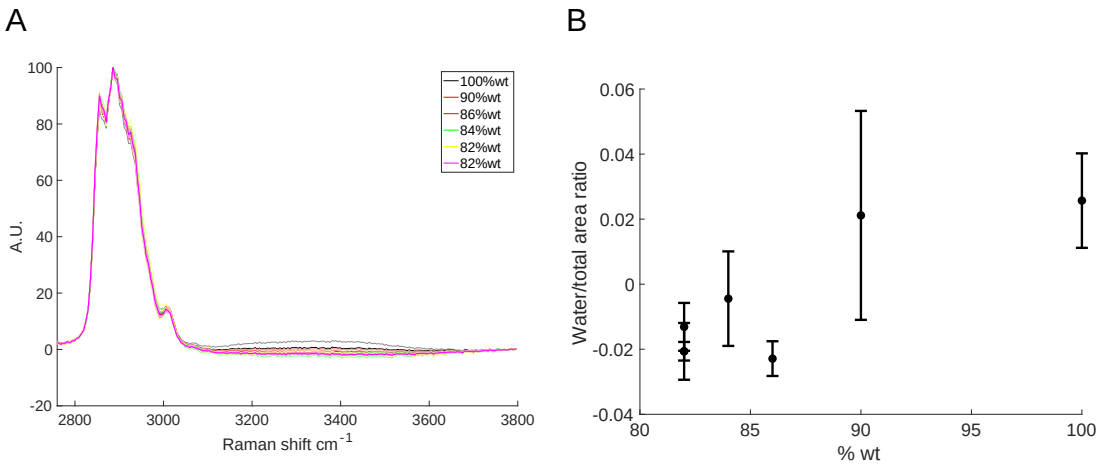


Figure 3-13 Figure showing Raman spectra in the HWN region of pork fat at different stages of dehydration and the relationship with water/total area ratio

A. Graph showing Raman spectra of pork fat specimens at 7 different stages of dehydration. Plotted lines are mean (n=5) for each %age change in weight normalised to CH stretch region between 2850 -2950 cm^{-1}, after pre-processing, shading in the same colour either side is +/- 1 SD.

B. Scatter graph of %age change in weight versus water/total area ratio. Points plotted are mean water/total area ratio (n=5) from spectra taken at each stage of dehydration, error bars +/- 1 SD.

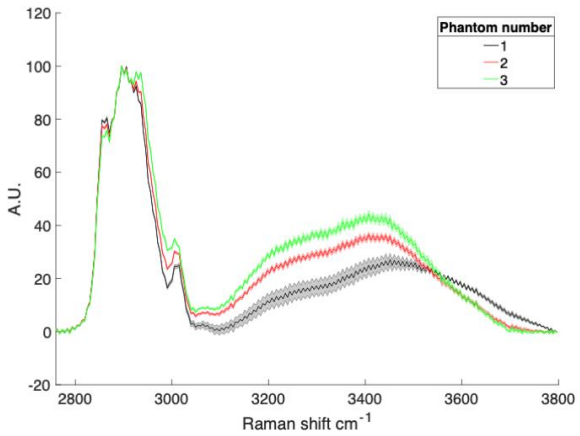
3.3.4 Mixed protein/lipid phantoms

Figure 3-14 demonstrates the results from 3 different mixed phantoms that had the same water concentration of 70%, but different protein/lipid ratios, and Figure 3-15 shows the results from 3 different mixed phantoms that had the same water concentration of 80%, but different protein/lipid ratios.

In both sets of phantoms it was found that with identical water concentrations, with a change in the lipid/protein environment the HWN spectra changes, and, significantly there is a change in the water/total area ratio. It can be seen that with a decrease in the relative concentration of lipid within the phantom, the AUC CH stretch region (relating to both protein and lipid peaks in these phantoms) decreases relative to the AUC water peak. It can also be seen that with the decrease in relative lipid concentration, there is a decrease in the AUC water peak – although the water concentration remains the same for each set of phantoms. These changes in the AUC CH stretch region and AUC water peak result in changes to the water/total area ratio – the mean water/total area ratio

increases significantly as the relative lipid concentration decreases in both the 70% water phantoms (displayed as mean water/total area ratio (SD); phantom 1- 0.4759 (SD 0.027) compared to phantom 2 -0.5486 (SD 0.0066); $P=4.5X10^{-5}$, and phantom 1 compared to phantom 3- 0.5796 (SD 0.0076); $P=1.2\ X10^{-6}$, and phantom 2 compared to phantom 3 P= 0.031) and in the 80% water phantoms (phantom 4- 0.6401 (SD 0.030) compared to phantom 5- 0.707 (SD 0.009); $P=2.3X10^{-4}$, and compared to phantom 6 - 0.7727 (SD 0.0047); $P = 2.1X10^{-7}$, and phantom 5 compared to phantom 6; $P=2.6X10^{-4}$) (one -way ANOVA).

A



B

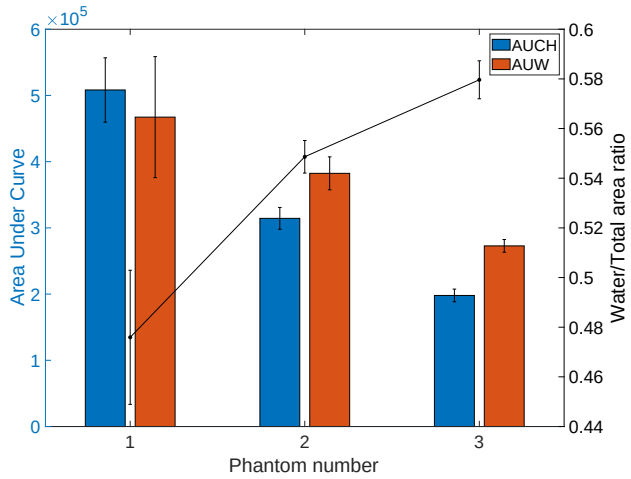
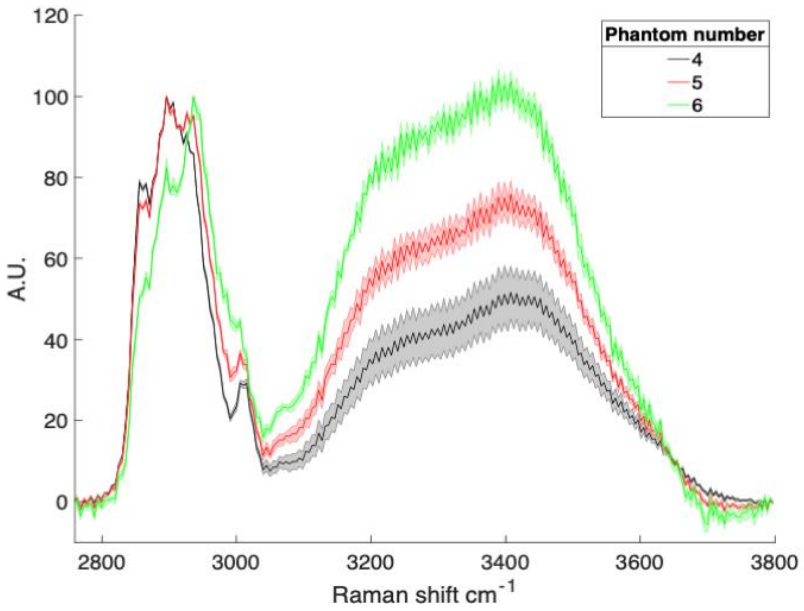


Figure 3-14 - Figures demonstrating the effect of a change in lipid/protein environment on the HWN spectrum in phantoms with a water concentration of 70%

Graph – A – baselined and normalised to CH stretch region between 2850 -2950 $cm^{-1}$for the phantoms 1 – 3, which all had a 70% water concentration but different protein/lipid ratio. B. Graph showing phantom number, Bars demonstrate the area under the curve of the CH stretch region (AUCH - Blue), and area under the curve of the water peak (AUW - orange), y axis on the left, with the water/total area ratio a black line with y axis on right. All values plotted are mean (n=5), error bars +/- SD. Water/total area ratio was significantly different between all phantoms (P< 0.05)

A



B

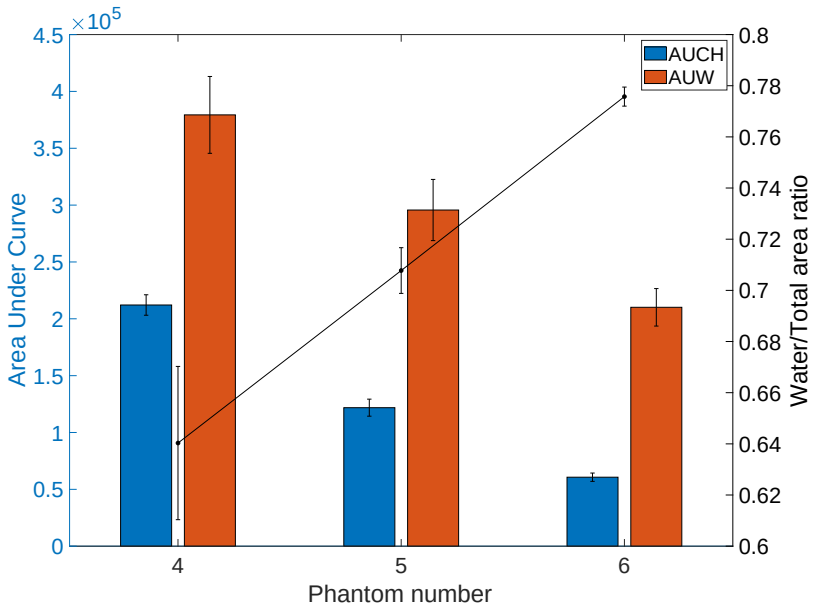


Figure 3-15 - Figures demonstrating the effect of a change in lipid/protein environment on the HWN spectrum in phantoms with a water concentration of 80%

Graph – A – baselined and normalised to CH stretch region between 2850 -2950 $cm^{-1}$for the phantoms 4 – 6, which all had an 80% water concentration but different protein/lipid ratio. B. Graph showing phantom number, Bars demonstrate the area under the curve of the CH stretch region (AUCH - Blue), and area under the curve of the water peak (AUW - orange), y axis on the left, with the water/total area ratio a black line with y axis on right. All values plotted are mean (n=5), error bars +/- SD. Water/total area ratio was significantly different between all phantoms (P< 0.001).

3.4 Discussion

3.4.1 NP3 can measure changes in water concentration at physiological concentrations

The results from the soya bean oil phantoms of varying concentrations demonstrate the ability of NP3 to detect changes in water concentration over a range from 30% - 95% water in a lipid only environment, with sensitivity to detect changes of 5%. The relationship between water/total area ratio and water concentration was linear, with an overall gradient of 1.24. However, there was some minor difference in the relationship between those phantoms with a higher water concentration (70-95%) compared to those with a lower water concentration (30-65%). The gradient of the lower water concentration line of best fit was 1.05, which suggests that the water/total area ratio is less sensitive to changes in the water concentration at lower water concentrations in lipid only environments. This could mean that over these broad ranges, that are likely to be encountered in human breast tissue, the relationship may not remain constant and this could affect the ability to quantify changes in water concentration. However, the difference between these two relationships are relatively small.
It should also be noted that the soya bean oil phantoms had no protein, which would be present in human breast tissue. The presence of protein (which has a CH peak close to the lipid CH peak) could affect the water/total area ratio and may affect these results, and the relationship between water/total area ratio.

The results from the serially dehydrated pork meat (a predominantly protein environment) demonstrates that there is a change in water/total area ratio with a presumed decrease in water concentration (as calculated by weight change). This demonstrates that NP3 is able to measure HWN spectra in biological systems, and it is able to detect changes in water concentration. The initial absolute water content of the pork could not be ascertained, but there was a loss of weight down to 45% of its starting weight over the course of its dehydration down to its ‘dry weight’, suggesting that water accounted for 55% of its initial weight which is a range encountered in human tissue [294]. The relationship between water/total area ratio and relative water concentration was not linear, the water/total area ratio changed little at higher relative water concentrations of 100 – 77% weight, but between relative water concentrations of 77-44% with a decrease in weight there was a decrease in water/total area ratio. This could

affect the ability of NP3 to quantify changes in water concentration in biological tissue, the reasons why this relationship may not be linear are subsequently discussed.

The results from the serially dehydrated pork fat demonstrated that the fat changes little in weight (by <20%), suggesting it had a very low water concentration initially of no greater than 20% which would be expected in hydrophobic tissue. It also dried out very quickly, with a number of measurements being the same, suggesting that it was truly dehydrated of any free water. There may be less bound water in fatty tissue as it is hydrophobic. As there is little bound water, it may explain why it dehydrates quickly and completely. There was negligible change in the water/total area in Raman spectrum over this range of 20%, and all values were close to zero (the reason for negative values is the fixed noise pattern oscillates around the zero baseline which give negative values for water AUC). The lowest soya bean oil phantom concentration was 30% water, and this fat dehydrated over a range of 20%, and probably has an initial concentration of < 30% water. It may be that in fat based environments in biological tissue of low water concentration of < 30% , the fat/total area ratio or any measure of HWN spectra, there is little change, and so changes in water content at low water concentrations (or high fat concentrations) cannot be detected. This suggests that for water concentrations <30% in predominantly lipid environments NP3 cannot detect changes in water concentration. These low water concentrations are likely to be encountered in some areas of fatty human breast tissue, and so this needs to be taken into account when analysing data from areas of low water concentration or high fat.

Summary

NP3 is capable of detecting changes in water concentration in soya bean oil phantoms, and these changes have a linear relationship with the water/total area ratio. The range that was investigated covers the physiological range likely of fat and water in human breast tissue. In biological tissue of serially dehydrated pork meat and fat the ability of the HWN region to detect changes in water concentration are less definitive. Although there was a trend of decreasing water/total area ratio with a decrease in water concentration in the protein rich environment of serially dehydrated pork meat, the relationship is not linear,

particularly when the pork meat had a higher water concentration. In pork fat, where the initial water concentration was very low, further dehydration did not lead to any measurable change in the water/total area ratio, suggesting that in fat environments of very low water concentration, HWN RS may not be capable of detecting changes in water concentration.

3.4.2 The effect of changes in water concentration on the HWN spectrum in three different environments

This section of the discussion aims to understand more about the spectra in the HWN region, the influences on the protein/lipid and water peak, and what needs to be taken into account when interpreting spectra and data from the HWN region to observe changes in water concentration or quantify changes in water concentration. Three systems were used to explore the effect of changes in water concentration on the HWN spectrum, protein only (gelatine/water phantoms), lipid only (soya bean oil/ water phantoms) and a biological system (serially dehydrated pork meat, mostly protein).

3.4.2.1 Protein only environment

In the protein only environment the spectral intensity of the protein peak increases, and the spectral intensity of the water peak decreases with a decrease in water concentration. This is expected as the relative protein concentration in the sampling volume is increasing with decreasing water. The AUC water peak decreases as water concentration decreases, and this relationship is linear. Similarly, the AUC protein peak increases as the protein concentration increases (and water concentration decreases), so the change in water/total area ratio seen with a change in water concentration is a result of these two spectral changes. This model is limited as the range of water concentrations is narrow (a 10% range), and so this relationship may not be true at lower water concentrations. However, it is in similar, simple, protein only systems that using HWN RS to quantify changes in water concentration has been validated previously. Masson *et al.* quantified water concentration changes in gelatin phantoms with a range of water concentrations from 66.67-90.91% [304] and Caspers *et al.* used a range of protein only solutions such as albumin, pepsin, lysozyme and urease to produce solutions of 60-80% water to calibrate their method of water quantification based

on water: protein ratios [295]. So, it may be that this relationship remains at a greater range of water concentrations than studied in this thesis.

This linear relationship between the protein and water peaks and water concentration in these simple protein models has also been used by those authors to suggest that this relationship is true in biological tissue, and therefore they use these simple systems (such as albumin and pepsin) to 'calibrate' their interpretation of the HWN spectra obtained in biological tissue (such as surgically excised tongue specimens) and provide a quantification of change of water concentration [174]. However, as has already been noted in the biological system of pork meat, this linear relationship is not necessarily true in more complex systems, which also involve the effect of sampling location, sampling volume and changing optical properties of the material with water concentration.

3.4.2.2 Lipid only environment

The results from the soya bean oil phantoms demonstrates that although in the lipid environment there is a linear relationship between water/total area ratio and change in water concentration the baselined, but not normalised, spectra shows that the water peak doesn't change much with different water concentrations, and it is the lipid peak that changes the most. This suggests that by using just the water/total area ratio, although validated in previous work, it may be an insufficient method to explain or understand the changes that occur in the HWN region with changes in water concentration.

In a lipid environment the relationship between change in water concentration and changes in the AUC are different compared to the protein system. The AUC water peak does not have a constant relationship with changes in water concentration, particularly at higher water concentrations, however the AUC lipid peak has a linear relationship of increasing with a decrease in water concentration (and increase in lipid concentration) with a small degree of variability throughout the range of water concentrations from 30 – 95%. There is a linear relationship between change in water concentration and the water/total area ratio in these soya bean phantoms. From the AUC graphs it can be assumed that the reason for this relationship is primarily due to changes in the AUC of the lipid peak, rather than changes in the AUC of the water peak.

This suggests that HWN RS in a lipid only environment is more sensitive to relative changes in lipid than changes in water concentration. In a phantom, such as the soya bean oil phantoms, where there is only lipid or water concentrations that are changed, the water/total area ratio remains constant with changes in water concentration, as if the lipid increases, the water concentration must also change. However, in a complex system which has protein, lipid and water and a number of different concentrations, the ability to detect and quantify changes in water concentration may be affected.

A reason for this observation is that the optical properties of the measured substrates will affect the pathlength of the laser photons and therefore the sampling volume, the pathlength also impacts on any differential self- absorption of the solution, as shown by the water absorption spectrum in the HWN wavenumber region. A limitation of the soya bean oil phantoms is that these emulsions are highly scattering. Intralipid, which is a similar emulsion composed of soya bean oil, has been used as a scattering agent to mimic the optical properties of breast tissue[308, 311]. So, it could be these changes in scattering properties of the soya bean oil phantoms that are causing the changes seen in the AUC peaks, as the changes in the highly scattering lipid may affect the changes in the water spectrum. However, creating a lipid only phantom over a broad range of physiological water concentrations is challenging, and a lipid only phantom that was capable of having a known water concentration, over a large range, that did not cause significant scattering, could not be found in the literature. Even if this is the cause for the changes seen, it is still relevant, as breast tissue is heterogenous and likely to contain areas of different scattering properties – indeed the reduced scattering coefficient (a measure of light diffusion in a tissue and reflection of scattered light from a tissue[314]) reported in the literature for breast tissue ranges from 8.3 to 31.8 cm^{-1} [346, 347] (one of the greatest ranges in biological tissues that has been reported [314]). Kim *et al.* found in skin that changes in water content of a sample led to changes in the optical properties of the sample [348]. For example, it was found that the reduced scattering coefficient (μ_s') doubled as the water content of the skin increased by 15%. This could be due to an increase in backscattered light with an increase in water content, and this was the main reason for variation in the intensity of Raman spectra in this study. If the

ability of HWN RS to quantify changes in water concentration is affected by changes in scattering properties, it must be considered when interpreting spectra.

UV-Vis measurements of the constituents of the soya bean oil phantoms were performed to investigate if the results were affected by the absorption characteristics of the phantoms' constituents. The results suggest that soya bean oil and Triton X100 had some degree of absorption, however it is at the same level throughout the relevant wavelength region and so it wouldn't disproportionately affect one part of the spectrum more than the other. Broadband absorption like this will change the entire HWN region equally, and so ratios between intensities are unchanged. However, the water absorption does vary throughout the spectrum, with an increase absorption between 1011-1050nm, which corresponds to a Raman shift of 2850 cm^{-1} and 3300 cm^{-1} at 785nm excitation. This may mean the water absorption would reduce the CH_3 stretch region of lipid disproportionately with increases in water concentration. This could affect the water/total area ratio as changes in water concentration are not affecting all of the spectrum to the same degree.

The absorption characteristics in skin were investigated by Kim *et al.* who found that there was a decrease in the absorption coefficient with an increase in water content in porcine skin [348]. This correlates with our findings that optical absorption properties change with changes in water concentration in biological systems, and so these observations need to be considered when interpreting spectra in the HWN region when using a 785 nm laser excitation. These findings will need to be considered in analysing spectra from breast tissue.

So, while broadband absorption will change the entire HWN region equally (when using a 785 nm excitation laser), and so the ability of using ratios or peak intensities for quantification of water content with HWN region remains unchanged, the absorption properties of water may mean that differing water concentrations may affect the HWN region spectra, and the relationship between water/total area ratio and changes in water concentration.

Summary

These results demonstrate that although the water/total area ratio changes with a linear relationship with a change in water concentration in a lipid only environment, the change in water peak doesn't have a consistent relationship with a change in water concentration, whereas the change in lipid peak shows a robust and consistent relationship with a change in water concentration. This suggests that HWN spectra of lipid rich tissue are more affected by changes in lipid concentration than changes in water concentration.

It has also been discussed that spectra in the HWN region are affected by the optical properties of scattering and absorption of the tissue measured, and that the scattering properties of fat, and the absorption properties of water will change with their respective concentrations, which may affect interpretation of HWN spectra in the assessment of changes in water concentration when using a 785 nm excitation laser. These findings are highly relevant to measuring in biological tissue such as breast tissue, as tissues are heterogenous with different concentration of water and lipid, and thus will have different optical properties. It is difficult to experimentally separate out the optical properties of our substrates measured here (soya bean oil and water) from their concentration, as they are inherently linked, i.e. if the concentration of lipid is increased, the scattering properties will change, and these cannot be isolated easily. Changes in optical properties may affect the HWN spectra to the same degree as changes in water or fat concentration.

Therefore, when interpreting spectra from the HWN region using a 785 nm excitation laser it must be considered that the HWN region is affected more by changes in lipid than in water concentration and by changes in the optical properties of the specimen. If there are changes seen between tumour and non-tumour breast tissue in the HWN region, it must be considered that changes in lipid concentration and optical properties affect the HWN spectra, as well as water concentration. Therefore, any changes in the water peak of the HWN spectra should not be considered to be due to solely a change in water content. These findings will aid interpretation of spectra when measuring a variety of microenvironments within breast tissue with the NP3 system.

3.4.2.3 Biological system

The results from the serially dehydrated pork show that, although there is a relationship between water change and water/total area ratio in pork meat, it is not linear as it is with the more simple protein model of gelatine phantoms at different concentrations. The pork meat is protein rich, however the results from the changes in the HWN spectra with water concentration are more similar to the lipid only phantom. The AUC CH stretch region peak increases with a decrease in water concentration. It is interesting that the protein peak intensity increases with a decrease in water content, even as we presume the amount of protein or lipid is unchanged as they are the same pieces of meat. A presumption in these experiments is that the weight loss of the pork is due to dehydration, and loss only of water. The meat only dehydrated for 24 hours in a laboratory, so it is unlikely that any decay or consumption of the protein occurred, and so we also assume that the amount of protein remained stable throughout the experiments. The AUC water peak does not have a consistent relationship with a decrease in water concentration, there is however a trend of increasing AUC water with a decrease in water concentration, although the meat lost >50% of its mass, which is presumed to be water. Biological systems are inherently more complex than simple phantoms constructed in the laboratory, and there are a number of phenomena occurring that may help to interpret these findings.

As water evaporates from the dehydrating pork it may not occur evenly from the specimen. It can be presumed that surface dehydration occurs more rapidly, and so the percentage weight change of the whole sample may not reflect the water change at the surface, which is where the concurrent Raman measurement is taken. The sampling depth of the NP3 system is estimated to be 500 µm (>50% of collected Raman photons are obtained from within this distance) [164]. Therefore, the majority of the collected photons will be collected from the surface, which is drying more rapidly from the bulk of the specimen, and thus may be an unrepresentative reflection of the specimen's water content. Another consideration is that some Raman scattering contribution will be from beyond 500 µm, from areas of greater hydration, and so within a single measurement different depths and water concentrations are being sampled, and this may be responsible for some of the observed intra-specimen variation.

The sampling of the specimen may also differ – certain areas of the specimen may have different water concentrations, and at every time point, it was not possible to ensure that the same part of the specimen was measured, which could lead to varying results.

Dehydrating biological tissue leads to a change in their optical properties. Rylander *et al.* demonstrated in skin that as it dehydrated the scattering particles such as collagen and organelles become more densely packed[349]. The pork specimen in these experiments visibly decreased in size, so as water leaves the specimen protein molecules come closer together increasing the amount of protein (and likely scattering) sampled in a given measurement. The specimens were also exposed to light, which may degrade or change biological pigments such as haem or carotene which may change the optical properties of the pork through its time of dehydration. These considerations may explain the greater variation in measurements with larger standard deviations compared to those of the gelatine or soya bean oil phantoms.

Another reason for the changes in HWN Raman spectra having a non-linear relationship with changes in water concentration in biological tissue is that water is sequestered in a number of different environments in biological tissues.
In the spectra of pork meat at its 'dry' weight (the point at which no further dehydration occurred) there is still a small spectral peak at 3280 cm^{-1} – this is due to the vibration of the NH group present in protein, which was also observed by Yang *et al.* [345] in dehydrated chicken, rather than signifying any residual water. It has been shown that water associated with presence of this NH group is more tightly bound [334]. Therefore the presence of the NH group in protein not only effects the HWN spectra because its vibration contributes a peak within the OH region, but also because water may be sequestered differently in the presence of the NH peak, which can affect the OH vibration, and subsequently the morphology, and behaviour of the HWN OH peak[350].

Unal *et al.* demonstrated that in biological tissue (bone), water is lost at different rates according to whether it is bound or unbound [351]. In their experiments they dehydrated bone of unbound water (by air drying) and bound water (by ethanol dehydration) and observed the changes in Raman peak intensities associated

with these two types of water, and found that unbound water loss followed a similar non-linear relationship to the pork meat dehydration, whereas the bound water loss followed a linear relationship with the associated Raman peaks. It is likely the water loss in the pork meat also represents unbound water. It may also be that as the specimen slowly dehydrates, water moves between compartments from bound to unbound, which would also affect the relationship of overall water concentration of the specimen with Raman spectra.

Figure 3-12 demonstrates that there are changes in the ratio of bound to unbound water, as pork meat dehydrates. As the water concentration of the pork decreases, the proportion of unbound water decreases and the unbound/bound water ratio changes from 1 (100% weight) to 0.43 (at 31% weight), this suggests that water losses are disproportionately from unbound water rather than bound water. These relative changes between bound and unbound water may contribute to our understanding as to why the water/total area ratio is not linearly related to a change in water concentration. The relationship between water concentration and the Raman peak intensity is dependent on the water type - the bound water peak Raman intensity has a linear relationship with changes in water concentration and unbound water Raman peak intensity has a non linear relationship water. This suggests that for the same number of water molecules, the Raman scattering cross section is dependent on the environment of the water. This means the AUC water peak is dependent not only on the total water concentration, but also what environment the water is in. The different water types are contributing differing proportions towards the overall AUC water peak at the different water concentrations which may explain why neither the AUC water nor the water/total area ratio has a linear relationship with water concentration in serially dehydrated pork.

This finding also helps to understand why the relationship between water/total area ratio and water concentration is different between the protein only environment of gelatine phantoms and the protein rich environment of pork. In pork dehydration, there are changes in where water is sequestered over time, whereas in gelatine phantoms that are composed of a set concentration this does not change (the unbound/water ratio in gelatine phantoms is 1.04 and does not change between 95% water and 85% water phantoms, data not shown).

The rationale for using the water/total area ratio area rather than specific parts of the water peak for calculation in water content is that it will include the changes between these types of water, whereas if only a narrow part of the water peak (e.g. just detecting unbound water) is used it may not detect changes in the other part of the water peak (e.g. changes in unbound water). A benefit of using a narrow part of the spectrum may be that the changes in water concentration with that part of the spectrum are more fully understood, e.g. if the bound part is used, one could assume a linear relationship with changes in water concentration. However, this would not give a full representation of what is occurring to the specimen in terms of total water concentration. It is not known whether the changes in water content between tumour and non- tumour breast tissue occur mainly in bound or unbound water, and so the whole water peak area should be used to capture any potential differences between tissue types.

Summary

These results demonstrate that in a protein rich biological tissue there is a relationship between water/total area ratio with change in water concentration, but it is not a linear relationship. It is also seen that the AUC water peak does not change in relation to a decrease in water concentration. The reasons for this relationship may be due to sampling technique, changes in sample structure and optical properties or changes in bound versus unbound water as the pork meat dehydrates.

Although these may be seen as ‘limitations’ of using serially dehydrated pork as a model, they represent challenges that are likely to be faced in biological tissue such as breast tissue. These properties that affect the HWN spectra are unique to each biological tissue, so what is found in pork tissue may not hold true of breast tissue. The shifts in water compartments are likely to be different between a tissue that is dehydrated to a certain water concentration, and a tissue that is in homeostasis at a certain water concentration. However, a highly relevant finding from these results is that there are changes to the HWN spectra according to where the water is sequestered, and changes in where the water is sequestered affects the water peak in a different manner. When measuring a biological tissue with a number of microenvironments, these are likely to hold

water in a different manner. Measuring the entire water peak therefore may give the best indicator of changes, but it may be there is a non- linear relationship, as seen in this biological pork tissue.

3.4.3 The water/total area ratio relationship with water concentration is dependent on protein/lipid environment

This section will discuss the influence on the HWN spectrum of changes to the protein/lipid environment on the ability to detect changes in water concentration, examining the results from the mixed gelatine phantoms.

In 3.4.2, the results suggested that the changes in the HWN region, and more specifically the water/total area ratio, is more a product of changes in lipid/protein concentration than of water concentration. It could therefore be hypothesised that the spectral changes with water concentration and the water/total area ratio will be dependent on the protein and lipid environment.

Caspers *et al.* used a range of different protein substrates (BSA, pepsin, lysozyme and urease) at a range of concentrations, to create a calibration tool for quantifying changes in water concentration using the HWN region [295]. There was a 15% difference in the proportionality constant calculated from the four different calibration proteins that was used to determine the changes in the water to protein ratio. This suggests that even between different simple protein/water mixes there is a different relationship between change in water content and the HWN spectra, and this may impact on the ability to assess changes in water concentration.

The water/total area ratio relationship between soya bean oil and gelatin phantoms over the same water fraction range is shown in Figure 3-16:

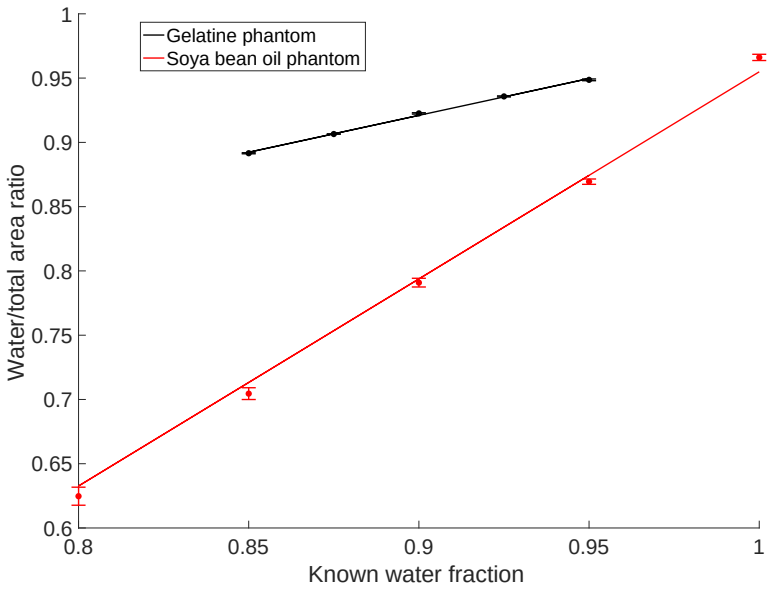


Figure 3-16 Graph showing linear relationship between known water fraction and water/total area ratio

Gelatine phantoms (gradient = 0.57, RMSE – $9X10^{-4}$) and Soya bean oil phantoms (1.61, RMSE = $4X10^{-16}$)

In both gelatine phantoms (protein only) and soya bean oil phantoms (lipid only) there is a straight line fit relationship –the changes in water concentration are proportional to the changes in water/total area. However, they have different gradients, that is the rate of change in the water/total area ratio is dependent on whether the water is in a protein or lipid environment. This suggests that changes between protein and lipid environments changes the relationship between water/total area ratio. A water/total area ratio from a protein rich environment will not equate to the same water/total area ratio in a lipid rich environment of the same water concentration.

The results from the mixed phantoms in Figure 3-14 and Figure 3-15 show that with a constant water concentration, the AUC of the water peak and the water/total area ratio changes with a change in protein/lipid ratio. This demonstrates that the water peak and the relationship of the water/total area ratio and water concentration changes is dependent on the protein/lipid environment.

These results also show that the lipid signal is dominant in the HWN region. Observing the Raman spectra from these phantoms demonstrate that even in phantom 5, which had the same concentrations of lipid (1%) and protein (1%),

the lipid signal dominates. What is also observed is that as the concentration of lipid decreases, the total AUC of both CH stretch region and water peaks decreases. This may be due to the scattering effect of the soya bean oil emulsion, meaning the Raman cross section of the entire phantom is reduced. However, also note that as the lipid content decreases, the water/total area ratio increases, this is because the AUC water peak decreases at a lesser rate than the AUC CH stretch region.

If the relative change in lipid concentration in the phantoms is calculated, it is equal to the relative change in the water/total area ratio that would be found in the change in lipid concentration in a soya bean oil phantom. For example, in the phantoms 4 – 6, the total oil concentration decreases by a total of 10% (from 15% lipid to 5% lipid), over this range the water/total area ratio increases by 14 % (from 0.69 to 0.83). Figure 3-16 B demonstrates the relationship between soya bean oil phantoms and water concentration has a gradient of 1.24, so in soya bean oil phantoms, if the water concentration increased by 10% (and therefore the lipid content decreased by the same amount), a 12.4% increase in water/total area ratio would be expected. This shows that with a change in lipid concentration, but with the same water concentration, there are changes in both the spectral features and water/total area ratio, that are comparable to if there were significant changes in water concentration.

Within this range of phantoms there are changes in protein concentration, however it is changes in lipid concentration that are predominately observed. This suggests that HWN spectrum may not identify more subtle changes in protein concentrations in the presence of lipids.

Summary

These results suggest that in a protein/ lipid environment, the HWN region mainly observes changes in lipids, more than changes in protein or water. A change in lipid concentration within a constant water concentration will change the water/total area ratio to a similar degree as a change in water concentration.

This suggests that using water/total area ratio within different protein/lipid environments would not accurately quantify changes in water concentration, as

altering the protein/lipid environment affects the water/total area ratio as much as a change in water concentration.

3.5 Conclusion

This chapter has demonstrated the influences on spectral measurements in the HWN region, and the ability to detect and quantify changes in water concentration. The results demonstrate the ability of NP3 to measure and quantify changes in water concentration in a lipid only environment at a range of physiological concentrations, and the relationship between changes in water concentration and the water/total area ratio in a biological tissue. However, the subsequent data analysis of measuring changes in AUC of the spectral peaks, and analysis of mixed phantoms with a constant water concentration and different protein/lipid environments demonstrate that the relationship between water/total area ratio is complex. The signal intensity of the water peak does not always increase or decrease with corresponding changes in water concentration, and the lipid peak appears to dominate the spectral features in the HWN region.

The rationale for performing these experiments was to understand the influences on spectra in the HWN region and the changes that occur with water concentrations. The ability to quantify changes in water concentration would be useful, as it would allow findings to be put in context with imaging modalities other than Raman spectroscopy. The findings from these experiments suggests that HWN Raman can quantify changes in water concentration, if the changes occur within the same protein/lipid environment. It is also possible to measure these changes in a number of different protein and lipid environments. However, the relationship between changes in water concentration and changes in the HWN region are dependent on the protein/lipid environment and so one cannot equate spectra or measurements of water concentration from one environment to another.

Previous studies have used the protein/water ratio from a set of protein only solutions to calibrate the system and then used these figures to quantify changes in water concentration in a different environment of biological tissue. In a protein only biological environment, such as skin [295] or corneas [296], this may be valid as the calibration tool is a protein only environment, so the error caused by changing

environments may be acceptable. However, the use of a protein only calibration tool for water quantification of spectra obtained in biological tissue including protein and lipid spectra as was done by Barosso *et al.* [174], based on our data, is likely to lead to significant inaccuracies in estimation of water content.

As biological tissue is composed of multiple different protein/lipid environments, of different optical properties, and water is sequestered in different ways, this complexity cannot be replicated to a sufficient degree to allow accurate calibration. The only reliable method of accurately quantifying water concentration within tissue is to use that tissue itself to provide a calibration tool. One method is by calculating water fraction by freeze drying breast tissue, taking a wet and a dry weight, and using a PLS model to correlate with the spectra obtained from different known water concentrations, as performed in brain tissue by Wolthuis *et al.* [294]. A limitation with this method is they were only able to produce a narrow range of water concentrations of 75-95%. Another method would be with serially dehydrating breast tissue and taking Raman spectra, similar to the serially dehydrated pork experiments. Access to appropriate samples obviously may limit the possibility of being able to perform these calibration experiments. Additionally, when analysing an area of breast tissue, measurements would be taken from a range of different lipid, protein and water environments, with constant changes within the specimen between these different environments. Therefore, calibration to enable quantification water content would require separate measurements for each environment likely to be encountered.

It may be that quantification of water content is difficult with HWN Raman in specific biological tissues due to limitations of accessing the appropriate biological tissue for experimentation and calibration. However, the HWN region has been shown to be able to detect changes in the HWN region between different protein/lipid environments, and can provide indicators of a high water concentration environment and a low water concentration environment. It can detect changes between lipid and protein environments and give individual information on where water is sequestered. Our hypothesis outlined in the introduction is that the spectral differences between cancerous and normal breast tissue is that tumour is a predominantly protein environment with a high water

content, and normal breast tissue is predominantly a lipid environment with a low water content. The experiments described in this chapter suggest that High wavenumber Raman Spectroscopy and the NP3 system is capable of detecting these differences, and therefore may be able to differentiate between cancerous and normal breast tissue necessary for Intraoperative Margin Analysis.

CHAPTER 4:
High wavenumber Raman spectroscopy for the identification of breast cancer in Fresh Frozen Samples

4.1 Introduction

In the background section to this thesis, the rationale for using HWN Raman spectroscopy for intraoperative margin analysis (IMA), and in particular the use of changes in water content to differentiate between tumour and non-tumour tissue was outlined. The major advantages of using the HWN region for diagnosis is that the acquisition times can be quicker and because of the reduced spectral features, analysis can be simplified, allowing a rapid assessment of large surface areas suitable for clinical IMA. Chapter 2 described the development of the NP3 system, which demonstrated the theoretical capability of measuring HWN Raman spectra in breast tissue, within the clinical environment, and Chapter 3 outlined an understanding of the influences on the HWN spectra, and how spectral data can be interpreted in relation to changes in water content which may aid in developing diagnostic algorithms for HWN spectra.

The next step in developing the NP3 Raman spectroscopy tool capable of performing IMA is to assess the ability of the system to differentiate tumour from non-tumour tissue. Raman spectroscopy has been shown to be capable of diagnosing breast cancer with clinical breast samples, but not in a clinically relevant timeframe for IMA [185, 262]. In laboratory conditions there has been assessment of the HWN region spectral differences between tumour and non-tumour human breast tissue, mainly observing changes in lipid and protein abundance [297]. There is evidence that there are some stark differences in lipid in the HWN spectrum between tumour and non-tumour tissue [271, 297], and there is potential for these differences to be used along with spectral information from the FP region to identify potential areas of tumour [299]. However, the differences in water content between tissue types and an assessment of using HWN RS alone

as a diagnostic tool in a clinical setting, with human breast tissue, has not been investigated. The aim of this chapter is to assess the diagnostic ability of HWN Raman to differentiate between tumour and non-tumour tissue in human breast specimens.

In the Raman literature there is an obvious focus on the differentiation between tumour and non-tumour tissue, as this is likely to be the most clinically relevant concerning IMA. However, there are a number of other potential uses for HWN Raman in breast cancer, particularly as the NP3 system is a needle probe which could be used for subcutaneous diagnosis. This system could be used for initial cancer diagnosis in addition, or in place of, core biopsy for rapid diagnosis, or information from these studies could be used in the development of a deep Raman system for screening of asymptomatic patients. Therefore, assessing the ability of NP3 system to provide a pathological diagnosis, or detect differences between pathological subtypes is an important part of developing the NP3 system for future clinical uses.

Assessment of NP3 thus far have been focused on the HWN region, as differences in water content may be key in rapid IMA. However, the NP3 system has a much broader spectral window due to the choice of grating that is required to capture the HWN region. Therefore, a large portion of the fingerprint region is also measured in any spectral acquisition. The fingerprint region is known to be able to differentiate between tumour and different types of non-tumour breast tissue [186], and it may give complementary and more specific biochemical information than the HWN region. The ability to capture both regions presents the opportunity to have a sensitive tool that can rapidly differentiate between tumour and non-tumour breast tissue with HWN spectral data and provide specific biochemical information from the fingerprint region that could give detailed pathological information.

An important part of differentiating between tumour and non-tumour specimens is an understanding of what is the ‘normal’ baseline against which the tumour tissue is being compared. When using changes in water content to differentiate between tumour and non-tumour, it must be considered that there are physiological variations in the water content of normal tissue between patients [306]

and that these physiological variations may be clinically relevant to disease states [280]. This will inevitably have an effect on the diagnostic ability to differentiate this normal tissue from tumour tissue, particularly as it may be that the extremes of the 'normal' physiological variation overlap with the pathological tumour state [276]. Therefore, the spectral differences observed between the normal specimens must be investigated as part of the pathological assessment.

Taking the aforementioned issues into account, the aim of this chapter is to assess the ability of the NP3 system to differentiate tumour from non-tumour tissue in human breast tissue samples. There is an emphasis on investigating the ability of detecting changes in water content to provide this diagnosis, and generally on the HWN spectral features that may provide diagnosis as part of IMA, but the spectral data acquired from the fingerprint region are also assessed to evaluate whether it can be used in conjunction, or in addition to, the HWN region for diagnosis or additional pathological information.

4.2 Methods

4.2.1 Human Tissue samples

Human tissue samples were acquired between 2011-2015 from the Royal Devon and Exeter Hospital. Freshly excised breast tissue removed from patients as part of their routine cancer treatments were processed following informed consent according to local protocols, having been approved by the Clinical Research Facility Tissue Bank steering committee (Ref: CRF320; Tissue bank ethics number 16/SC/0162; further information at https://exetercrfnihr.org/about/rde-tissue-bank). Inclusion criteria were females aged >16 and able to consent for the procedure, with a malignant tumour of greater than 2 cm determined by pre-operative examination or imaging. Routine demographic data were collected from clinical notes and linked anonymously with the research samples. The specimen was sliced open fresh, and a 3 mm core biopsy was taken from the tumour and a separate 3 mm core biopsy taken from surrounding normal tissue, distant from the tumour edge (to allow a matched pair of biopsies of tumour and normal breast tissue from each patient). This was performed by a histopathologist or pathology practitioner. The rest of the specimen then underwent routine histopathological analysis. The research core biopsies were labelled anonymously and

immediately frozen in liquid nitrogen and stored at -80 ºC. Subsequently, 5 µm sections were taken from all of the frozen biopsies for Haemolysin and Eosin (H+E) staining and underwent analysis by a histopathologist from Gloucestershire Royal Hospital for a histopathological diagnosis, and sections taken to be fixed on barium fluoride slides to enable microscopic Raman imaging. The remainder of the biopsy (which varied in volume between roughly 2 mm^3 to 5 mm^3) was available for needle probe analysis.

4.2.2 Needle Probe Raman Spectroscopy configuration

The Raman needle Probe was configured as outlined in Chapter 2 for the NP3 system. In brief, a 785 nm laser excitation was used, with a Kaiser Spectrometer with a broad grating with an InGaAs camera, with fibre optics within a needle probe for light delivery/collection.

Daily calibrations were performed with Neon Argon lamp, Ethanol, water, PTFE Paracetamol and aspirin. This was used to calibrate the Raman shift using the inbuilt calibration tool in the SOLIS software. A dark signal measurement of 5 second acquisition and 5 accumulations were taken daily for baseline subtraction (method detailed in Chapter 3).

4.2.3 Fresh Frozen sample Breast Tissue measurements

The fresh frozen biopsy samples were measured individually. The sample was removed from the freezer, thawed at room temperature and measured immediately to minimise loss of water by evaporation before being disposed of appropriately. The whole process took less than 5 minutes for each sample.

Measurements were performed at an ambient temperature of 21 – 24 ºC in a dark room. The specimen was removed from its container once thawed and placed on a calcium fluoride slide. The needle probe was then put in contact with the specimen for measurement. 5 – 8 different areas of the specimen were measured at random, depending on the size of the specimen. Measurements were taken as 5 second acquisitions and 5 accumulations in each of the 5-8 different areas.

4.2.4 Estimation of water content by dehydration

To obtain a correlation between the HWN Raman spectra and calculated water/total area ratio and water content of breast tissue, a series of dehydration experiments were performed. Five tumour specimens and two non-tumour

specimens were chosen for the experiments – they were the largest specimens available (to enable measurable change in weight).

Immediately after the initial Raman measurements for tumour diagnosis experiments, samples were weighed using weighing scales accurate to 0.0001 g. This initial weight was the 'wet weight'. The samples were then placed in the laboratory environment to dehydrate over time and were measured every hour until the weight was unchanged for two consecutive readings (approximately 6 hours), and this was recorded as the 'dry weight'. Weight loss was presumed to be wholly due to water dehydration. Water fraction was calculated as per Equation 4-1:

$$Water\ Fraction = \frac{Weight\ of\ specimen - 'dry\ weight'}{'wet\ weight'\ of\ specimen}$$

Equation 4-1 Calculation of water fraction of dehydration of specimens

Immediately prior to the weighing of the specimen at each time point, measurements were taken as per the other Needle probe Raman measurements. This resulted in a series of Raman spectra at each stage of dehydration for each specimen.

4.2.5 Data Processing

Spectra were recorded using SOLIS software (Andor,UK) and processed in Matlab (Mathworks, USA). For each measurement the mean of the 5 accumulations was taken. A smoothed dark noise baseline was obtained by a Savitzky-Golay filter with a 7th order polynomial of the dark noise reading and this was subtracted from the spectra. Spectra that had high levels of background fluorescence that resulted in little visible measurable Raman spectral data were rejected (n=3).

4.2.5.1 HWN region

For calculation and analysis of the water/total area ratio the HWN spectra underwent 3rd order polynomial baselining with no normalisation (as demonstrated in Figure 4-1 A). This is similar to other studies that have investigated water / protein / lipid ratios , where the data has undergone minimal

pre-processing with fluorescent baseline subtraction only prior to calculation of the ratio [174, 304].

Pre-processing for the High wavenumber (HWN) data (2600-3800 cm^{-1}) underwent fluorescent baseline subtraction with a 3rd order polynomial and minimum/maximum normalisation (whereby the lowest reading in the individual spectrum =0 and the highest reading = 1). 3rd order polynomial background subtraction has been used previously for HWN Raman in biological tissue [240, 304] and has been shown to be optimal for baselining this region compared to other polynomials[337]. Other methods of pre-processing of spectral data such as normalising using the median Area under the Curve, normalising using the mean Area under the Curve, and normalisation with minimum/maximum without baselining were trialled (shown in Figure 4-1). This sequence of pre-processing was used as this was the method that gave the best separation between samples.

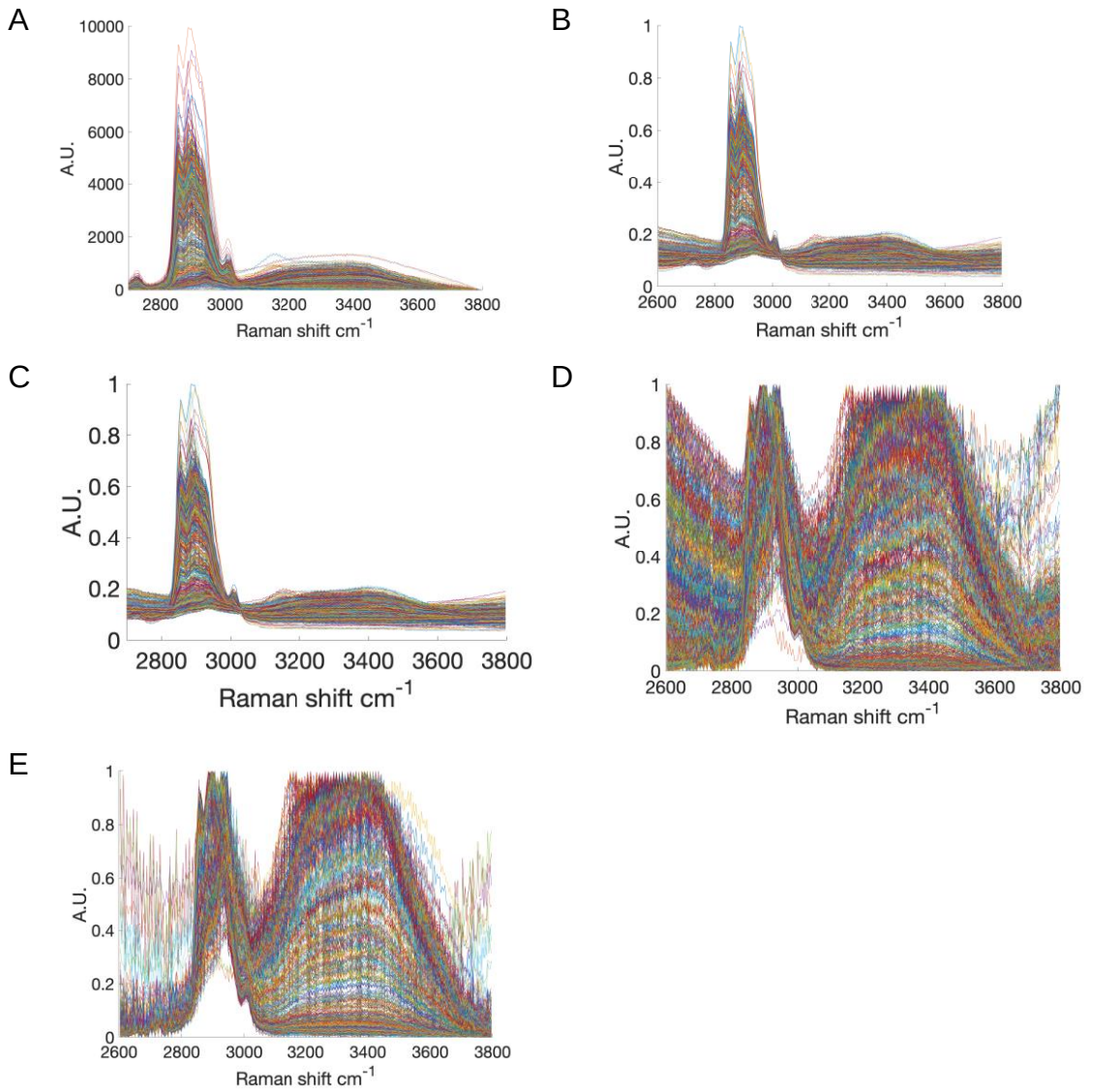


Figure 4-1 High wavenumber spectra of human breast tissue using different pre-processing methods A Fluorescent Baseline subtraction using 3rd order polynomial. This was used for the calculation of water/total area ratio. B. Normalisation of spectra using the Median Area Under the Curve C. Normalisation of spectra using the Mean Area Under the Curve D. Normalisation using Minimum/Maximum E. Fluorescent Baseline subtraction using 3rd order polynomial followed by normalisation using Minimum/Maximum. This is the method that was used for HWN spectral analysis.

4.2.5.2 Fingerprint Region

The Fingerprint (FP) region data (900-1886 cm^{-1}) had a more complex fluorescent background to consider, and needed to be pre-processed differently to the HWN region. Using different pre-processing for the different spectral regions of fingerprint and HWN is commonly performed when both are being measured [107, 304]. An example of a raw spectrum is displayed in Figure 4-2, with a hand-drawn line demonstrates the complex shape of the background that needs subtraction.

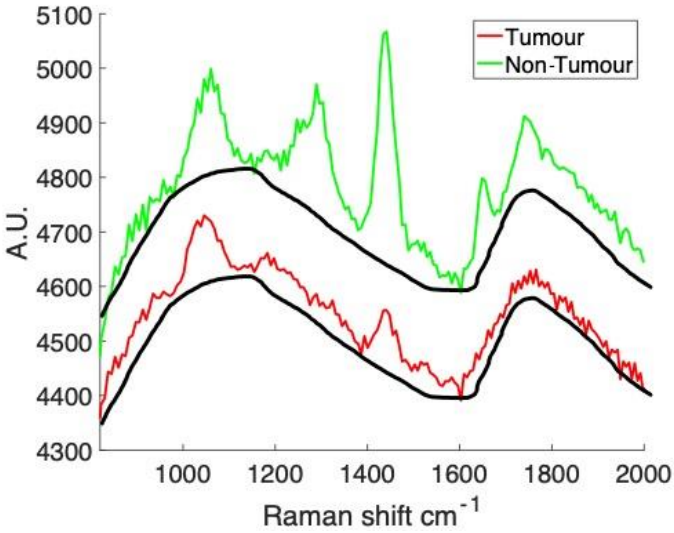


Figure 4-2 Demonstrating the complex fluorescent background baseline of breast tissue specimens Representative raw Raman spectrum from a tumour (red) and non-tumour specimen (green). Hand drawn 'baseline' (black) suggests the ideal baseline that would be fitted to the spectra to obtain optimal spectral pre-processing

Baseline subtraction is performed to subtract the fluorescence sufficiently to enhance spectral features to allow greater clarity for interpretation, however if the spectrum is 'overfitted' it can remove important spectral features and have the opposite effect on interpretation [100]. The process is a balance between underfitting, which leads to 'false' peaks of background being interpreted as Raman peaks, and overfitting, with Raman peaks being subtracted out along with the baseline.

Another method of processing Raman spectra, particularly for PCA and LDA analysis (which was one of our planned methods of analysis) is to perform the analysis on spectra that have been normalised (to provide standardisation in intensity between spectra and allow comparison) but without any baseline subtraction. This removes the human error of visual interpretation of a baseline and avoids under or over fitting the baseline to the spectrum. Compared to the HWN region, the spectral features of tumour and non-tumour breast tissue in the fingerprint region are very well documented. Therefore, we processed the Raman spectra without a baseline and trialled a number of different polynomial baseline subtractions (what is displayed is representative of the baselines attempted, rather than exhaustive) (Figure 4-3). A tumour spectrum was used, as this was the group that was the most difficult to fit and gave the most disparate results – non-tumour spectra were affected less by the different methods of baselining.

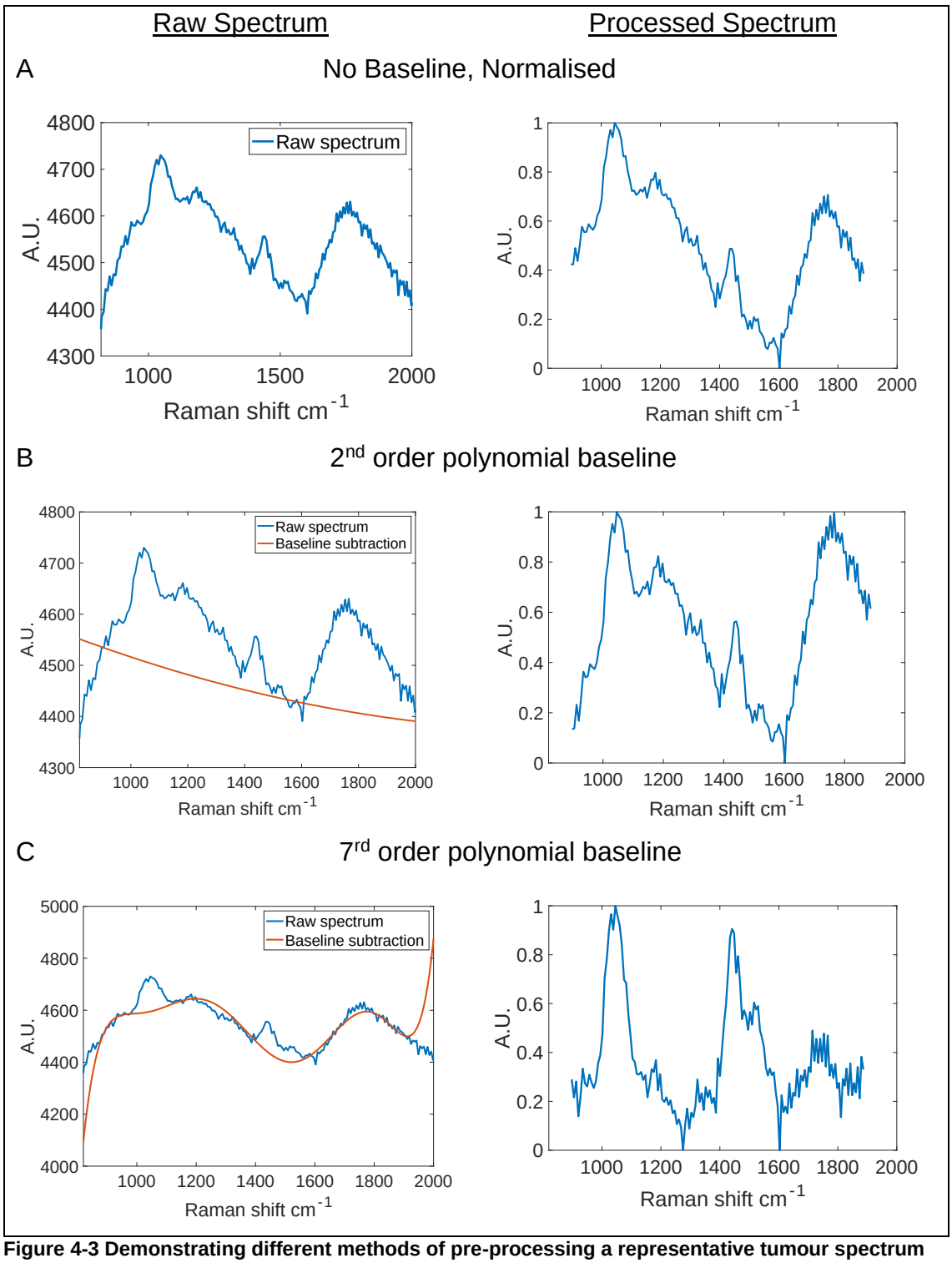


Figure 4-3 Demonstrating different methods of pre-processing a representative tumour spectrum

On the left are the raw spectra, with the polynomial baseline that is calculated for subtraction, on the right are the resultant spectra after baseline subtraction, and min/max normalisation. A. Spectrum does not undergo baselining, just normalisation. B. Spectrum undergoes 2nd order polynomial baselining and normalisation. An example of baseline underfitting, as there remains significant fluorescence baseline artefact. C. Spectrum undergoes 7th order polynomial baselining and normalisation. An example of baseline overfitting, although all fluorescence baseline is subtracted, some of the spectral features are also lost.

In the spectrum that did not undergo baselining and in the spectrum that has been 'underfitted' with a second order polynomial, it can be seen that interpretation of

peaks is difficult because the baseline is significant, there is no amide III region (1200-1300 cm^{-1}[1103]) which would be expected in protein rich tissue, and from the literature[352], and there is a large peak at 1700-1900 cm^{-1}. There are no Raman active molecules beyond 1800 cm^{-1}, so it is certain that this is an artefact of fluorescence. When we performed a PCA analysis of the spectra that had not been baselined, or baselined with a second order polynomial, the principle components were difficult to interpret due to the remaining baseline, and so made it impossible to analyse the biochemical differences that were separating the data.

The spectrum that has been processed with a 7^{th} order polynomial shows signs of overfitting. The amide III region, which would be expected to be present, is absent, and so some of the specific biochemical features for tissue differentiation are lost. However, it has successfully subtracted the fluorescence from 1700-1900 cm^{-1}.

We then used a 6^{th} order polynomial to fluorescence baseline subtract, which gave the optimal results (Figure 4-4). The processed spectrum shows clearly the amide III region, which was to be expected, and there is no significant peak in the 1700-1900 cm^{-1} region. When compared to a previous study's results, which were also obtained in fresh frozen breast tissue using a 785 nm laser excitation, the spectrum is similar[352]. A 6^{th} order polynomial has also been used in a previous publication for baseline subtraction in the fingerprint region when measuring fluorescent biological tissue [107]. This ability to perform specific biochemical analysis of the spectrum and the concordance with previous results suggested this was the optimal method for processing our fingerprint spectral data, and so this method was used in the rest of this study.

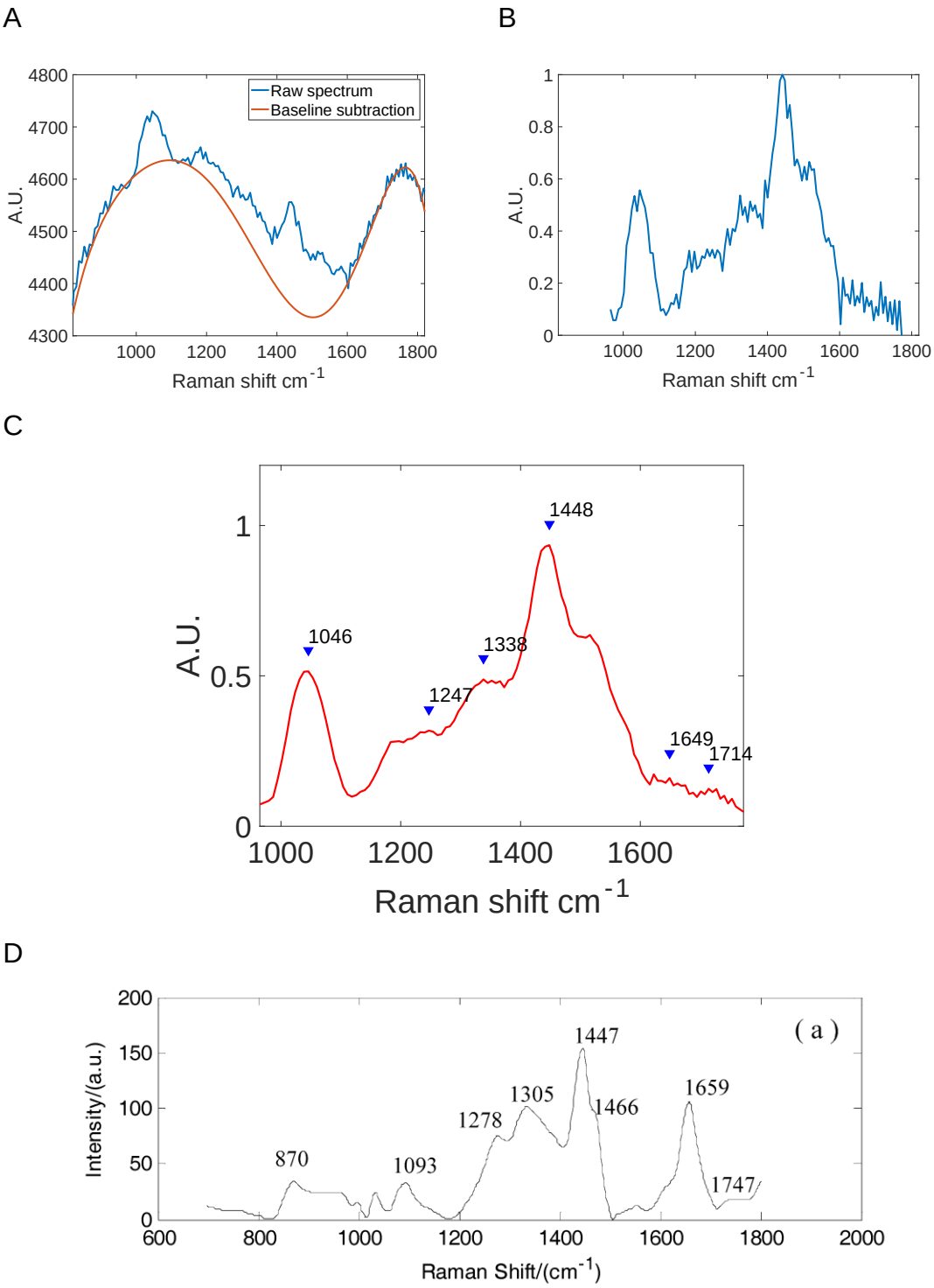


Figure 4-4 Demonstrating pre-processing tumour spectrum with a 6th order polynomial flourescence baseline subtraction prior to min/max normalisation.

A Representative raw tumour Raman spectrum, with the fitted baseline to be subtracted. B Representative tumour spectrum after pre-processing. C. Representative tumour spectrum (smoothed with a savitzky-golay span of 5, 1st order polynomial) with peaks highlighted with corresponding Raman shift number. D. Figure of Raman spectrum from the literature obtained with a 785 nm laser excitation demonstrating similarity in Raman peaks with our spectrum in C. Reproduced from reference Li *et al.* [352],(copyright the authors 2017, reproduced with a creative commons licence 4.0 (https://creativecommons.org/licenses/by/4.0/))

4.2.5.3 Concatenated data

Concatenation was then performed where the pre- processed fingerprint spectrum was 'stitched' with the high wavenumber data to allow analysis without the Raman inactive region of 1900-2600 cm^{-1} (as shown in Figure 4-5).

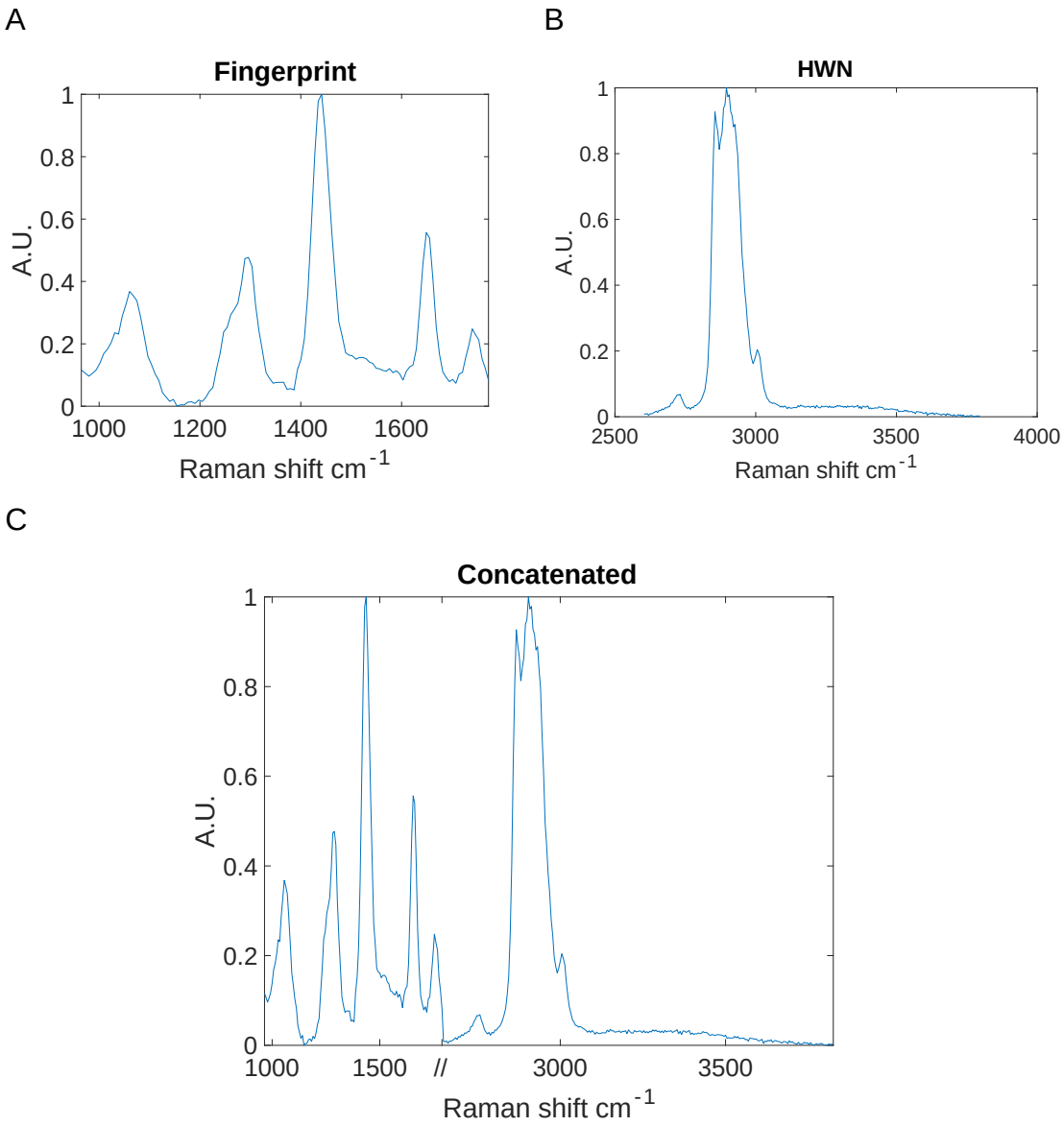


Figure 4-5 Representative baselined and normalised Raman spectrum obtained from a single non-tumour specimen, with demonstration of the different spectral regions measured

A Fingerprint region B. Highwavenumber (HWN) region C. Concatenated data composed of the HWN and Fingerprint region 'stitched' together

4.2.6 Data analysis

Spectra were analysed both as individual, independent spectra (e.g. each different area on a single sample viewed as a discrete spectrum and so each specimen had a number of spectral outputs that contributed to analysis), and as mean/median specimen spectra (e.g. each sample had a single combined spectral output).

The pre-processed HWN spectra were analysed using the water/total area ratio calculation using the protocol outlined in Chapter 2. The water/total area ratio was calculated for each individual spectrum and then the mean ratio from each specimen was used and categorised as coming from either a tumour or not tumour specimen. The standard deviation was calculated for mean data.

Binomial Logistic Regression was then used to obtain the optimal water/total area ratio threshold that could be used to predict if a spectrum was from a tumour or not tumour specimen. Classification tree machine learning was also used with a 2- way split as an alternative method to obtain the same goal. In data with more than two classifiers (e.g. carcinoma type and grade), a multiclass error-correcting output codes (ECOC) model using support vector machine binary learners was trained for prediction. For each model a Receiver Operating Characteristic (ROC) curve was plotted, with the y axis being true positive rate, and x axis 1-specificity (or false positive rate). The optimal threshold was calculated as the point at which the test gives the optimal accuracy and balance of sensitivity and specificity. The accuracy of these models can be assessed by the Area Under the Curve (AUC) of the ROC curve- a ROC with an AUC of 1 is a perfect test with 100% sensitivity and specificity, and therefore, the closer the AUC is to 1, the more accurate the model.

These techniques were then cross validated by a number of methods, using built in functions within MatLab. Different cross validation techniques were used to assess which were the most suitable to use for future data analysis. No technique is perfect, and there is no ‘correct’ one for use in a particular scenario; it is important to acknowledge the potential risk of bias due to each method, when deciding which method of cross validation to use [353].

'Hold out' cross validation is where 30% of the data is held out, the model uses the remaining 70% of data as the 'training set', and it then uses this output to predict the outcome of the previously held out 30% as the 'prediction set'. The advantage of this technique is that none of the training set data is used for the prediction set, meaning the prediction set is truly 'new' data to the algorithm. However, with a limited data set there is a possibility of bias, as the test set (which does not inform the algorithm) may contain important information (e.g. it may contain all the data from a sub-population)[354].

'K folds' cross validation is where the data is randomly split into the number of 'folds' stipulated, e.g. in K folds 5, the data is split into 5 groups. One group is left out, the remaining groups are used as a 'training set' and this output is used to predict the outcome of the previously left out group as the 'prediction set', this is then repeated, but with the next group left out etc. until all groups have been held out and used as a 'prediction set'. The advantage to this technique, particularly with a smaller data set, is that all the dataset is used in at least one iteration of the training set, reducing the risk that important information or a sub-population is not included in algorithm construction[354], however this increases the possibility of overestimation of the algorithm prediction accuracy [355].Leave-One-Out-Cross-Validation (LOOCV) is the logical extreme of a K folds cross validation, with the number of folds being n-1. LOOCV is the most exhaustive cross validation procedure and is commonly used in clinical studies providing a 'subject-wise' cross validation method (the training set does not include the prediction set data), and is considered to be particularly appropriate in the evaluation of diagnostic models [353, 355].

For each of the cross validation methods the accuracy of the test was assessed by the outcome in the form of a confusion matrix with the sensitivity (true positive rate, or true positive / true positive + false negative), specificity (true negative rate, or true negative / false positive + true negative) and overall accuracy of the analysis (true positive + true negative) / true positive +true negative + false positive + false negative) calculated.

The individual specimen data also underwent analysis with a voting threshold technique. The optimal water/total area ratio was used as a cut off to classify

each spectrum as either ‘tumour’ or ‘non-tumour’. The optimal water/total area ratio that had been determined from the binomial logistic regression model without cross validation was used as this used all available spectra, whereas in the cross validated models a certain proportion was held out, and so reduced the size of the training set. Algorithms were constructed in Matlab to classify the entire sample based on the proportion of ‘tumour’ or ‘non-tumour’ spectra for that sample. The rules of the algorithms for Voting threshold 1 or 2 are detailed in the Results section. This gave a single binary output for each specimen, as the entire specimen either being classified as tumour or non-tumour.

All spectral data (from individual or mean spectra from all spectral regions) was analysed by Principal Component Analysis (PCA). Principal components (PC) describe data as sections of spectra and the degree of variance between spectra. It therefore describes each spectrum as a score for each PC and P values assess the difference between the scores between spectra. It is an unsupervised test and the potential differences between groups (e.g. tumour or not tumour) are not acknowledged by the analysis. It is a commonly used technique in Raman spectroscopy to analyse differences in spectra.

Spectral data were mean centred and underwent PCA, with 12 separate component separations. Scores had P values calculated by one-way ANOVA and were regarded as significant if $P<0.01$, and this cut off was used to highlight the components that required plotting. The component scores were plotted on graphs with the ability to differentiate between different pathologies/demographics (e.g. tumour or non- tumour) by colour.

If there was separation on PCA by P score values, a Linear Discriminant Analysis (LDA) was performed with the PCA fed scores. LDA is a supervised test, where the differences in the samples (e.g. if they came from tumour or not tumour) are loaded into the test and used to classify the spectra and so the known pathological state is acknowledged in the analysis. In our analysis the pathology/demographic label was loaded with the scores, with 100 bins. This was then cross validated by a Leave-One-Out Cross Validation (LOOCV) technique, where a sample is left out, the PCA fed LDA is performed on the rest of the data

to have a ‘training’ set, and this is then used to predict the outcome of the sample that was left out. This is then performed n-1 times.

Spectral data was analysed according to patient/tumour subtypes. Tumour spectra was classified according pathological subtypes, and non- tumour spectra classified according to patient physiological factors of Menopausal status (pre and post) and age. Age is a continuous variable, whereas the analysis to be performed relied on classification according to a binary or ordinal variable. To split age into a number of different age brackets would result in a small number of patients being in each group and under- powered calculations. Therefore, age was divided into a binary variable, with a cut off of either 60 (i.e. below or above 60), or cut off of 50 (i.e. below or above 50). This lower age bracket was used as there is good evidence that breast composition with regards to water content is markedly different in much younger age groups regardless of menopausal status[280], and if the cut off had been lower (e.g. 40), there would have been too few patients in the group to allow analysis. In these analyses, if the subtype had not been collected (left blank on the patient information spreadsheet), the spectra was excluded from that analysis.

4.2.7 Raman micro spectrometry experiments

To enable matching between the measured Raman spectra obtained from the needle probe and histological features of the breast tissue, a series of experiments were conducted on a micro spectroscopy system. The results from the fresh frozen specimen were examined to determine which required corresponding microscopic analysis. Two specimens were chosen to have microscopic analysis as they had some Raman spectra indicating ‘non-tumour’ tissue in the tumour specimen, and the prepared slides were all of sufficient quality, with an adequate amount of tissue, to allow detailed analysis.

H+E stained slides were digitised using an upright light microscope (Carl Zeiss AG,Oberkochen, Germany) at a X 20 magnification (20X/0.30 objective, Leitz Wezlar, Germany), with a CMOS camera (Model - acA1920-155uc, Basler AG, Ahrensburg, Germany) and images were captured using Manual MSI, Microvisioneer software (Freising, Germany). The digitised images were printed, and the original slides were analysed by a second consultant histopathologist

from the Royal Devon and Exeter Hospital for a more detailed microscopic diagnosis at X 40 magnification, and pathological features of the slide, which was annotated on the digitised image.

The corresponding tissue slices that had been prepared on BaF_2 slides were thawed at room temperature and immediately measured. The Renishaw InVia confocal spectrometer system (Renishaw, UK) was used for Raman measurements after calibration of the system with silicon, green glass, PTFE and Neon Argon lamp. A slide scan was then performed with white light, at X5 magnification and this was correlated with the annotated digitised H+E slide. Areas of interest had Raman measurements taken using a NPlan objective (Leica, Germany) with magnification X50. An excitation source of a 785 nm laser was used with an output of 300 mW to collect spectra in the wavelength range of 2100 – 2750 cm^{-1}, with a single accumulation and 5 second acquisition time. Raman spectra was analysed in Matlab, and had pre-processing by fluorescent baseline subtraction by a first order polynomial only.

4.3 Results

4.3.1 Patient/ Tumour demographics

Fresh frozen samples from 96 patients were used for this study. The patient and tumour characteristics of the patients are shown in Table 4-1.

DEMOGRAPHIC DETAIL	NUMBER (%) / MEAN (± SD)
PATIENT DEMOGRAPHIC	
MEAN AGE	67.3 (±14)
POST MENOPAUSAL	80 (83%)
POSITIVE FAMILY HISTORY	24 (25%)
ROUTE OF REFERRAL - SYMPTOMATIC	85 (89%)
TYPE OF SURGERY	
SITE OF OPERATION - LEFT BREAST	57 (59%)
WIDE LOCAL EXCISION	50 (52%)
MASTECTOMY	46 (48%)
SENTINEL NODE BIOPSY (OF 95 THAT HAD AXILLARY PROCEDURE)	73 (77%)
AXILLARY NODE CLEARANCE	23 (23%)
TUMOUR CHARACTERISTICS	
DCIS ONLY	3 (3%)
INVASIVE CARCINOMA	93 (97%)
MEAN SIZE OF TUMOUR (MM)	30.2 (± 11.8)
INVASIVE TUMOUR CHARACTERISTICS (N=93)	
POSITIVE LYMPH NODES	37 (39%)
MEAN NUMBER OF POSITIVE LYMPH NODES	1.3 (± 2.3)
INVASIVE ASSOCIATED WITH DCIS	72 (77%)
INVASIVE CARCINOMA GRADE (N=93)	
- GRADE 1	2 (2%)
- GRADE 2	54 (58%)
- GRADE 3	37 (39%)
CARCINOMA TYPE (N=93)	
- DUCTAL	64 (67%)
- LOBULAR	21 (23%)
- MIXED	5 (5%)
- MUCINOUS	3 (3%)
RECEPTOR STATUS (N= 96)	
ER +	82 (85%)
HER2 +	17 (18%)

Table 4-1 Table of patient and tumour characteristics for the samples that were measured in this study.

Note that in demographics with a dichotomous outcome (e.g. post- menopausal) only one is included, and it can be presumed the other patients were in the other category (e.g. pre-menopausal). ER = Estrogen receptor, HER2 = Human Epidermal Receptor

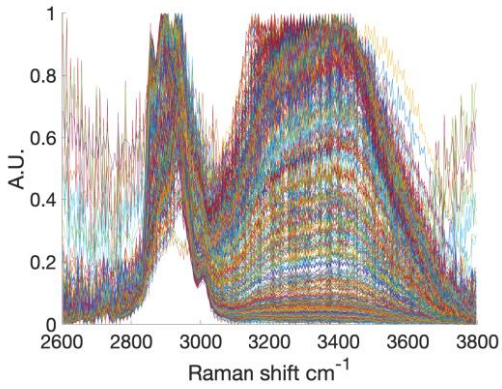
4.3.2 Fresh frozen specimen measurements

4.3.2.1 Spectra obtained

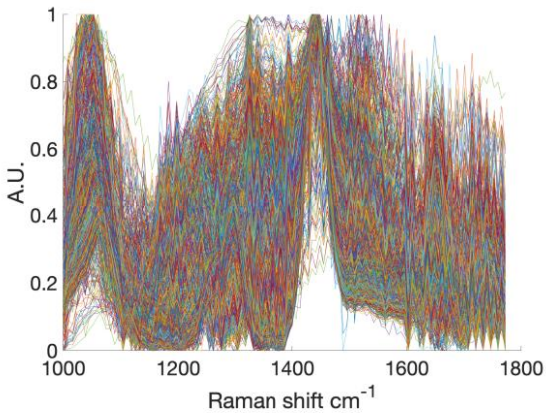
In total 192 specimens were measured (96 tumour specimens, 96 non-tumour specimens), with a total of 1335 spectra obtained.

After pre-processing the data of all spectra obtained is shown in Figure 4-6.

A



B



C

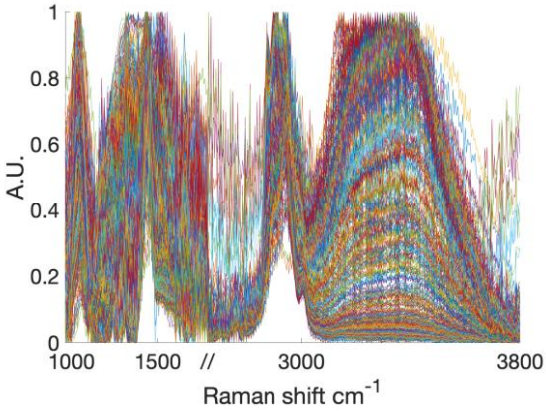


Figure 4-6 Graphs showing spectral data obtained from all measurements (n=1335) after pre-processing

A. HWN region B. FP region C. Concatenated data. In all graphs the spectra shown have had fluorescent baseline subtraction and min/max normalisation.

4.3.2.2 Water content analysis to diagnose breast cancer

After pre-processing, it is evident that there are visual differences in the HWN region between spectra obtained from tumour specimens and non-tumour specimens. Figure 4-7 shows there is a clear difference in the water peak between tumour and non-tumour specimens, with tumour specimens having a larger water peak than non-tumour specimens, when spectra have been normalised to the CH_2 stretch region (2935 cm^{-1}). Non-tumour specimens also have a predominately lipid peak at the CH stretch region (2800-3040 cm^{-1}), whereas tumour specimens have a predominately protein peak at the CH stretch region.

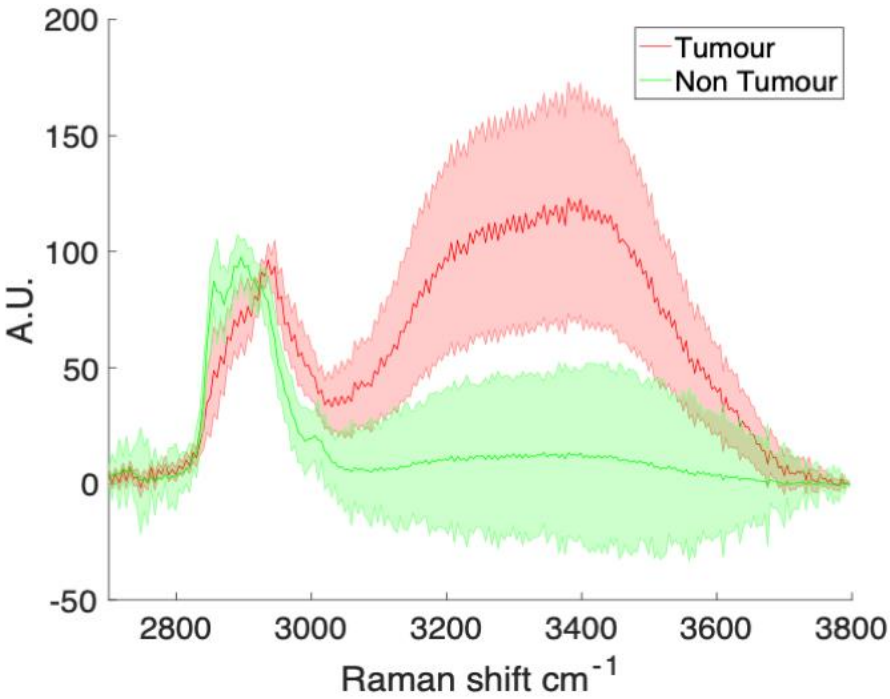


Figure 4-7 Mean HWN spectra of all specimens (n=1335) according to diagnosis demonstrating changes in water content between tumour and non-tumour specimen spectra

Data is mean of spectra from all tumour specimen spectra (n=672) and all non-tumour specimen spectra (n=663), fluorescent baseline subtraction with a 3rd order polynomial, and normalised to the CH peak at 2935cm^{-1}. Shading either side of solid line in same colour is +/- 1 S.D. Red spectra=measurement from tumour specimen, Green spectra = measurement from non-tumour specimen

The water/total area ratio was calculated for each spectrum, and the mean and median of the ratio calculated for each specimen. Figure 4-8 shows that there is a significant difference for both the mean ($P=1.6X10^{-42}$) and median ($P=2.3X10^{-30}$) water/total area ratio between tumour and non-tumour specimens.

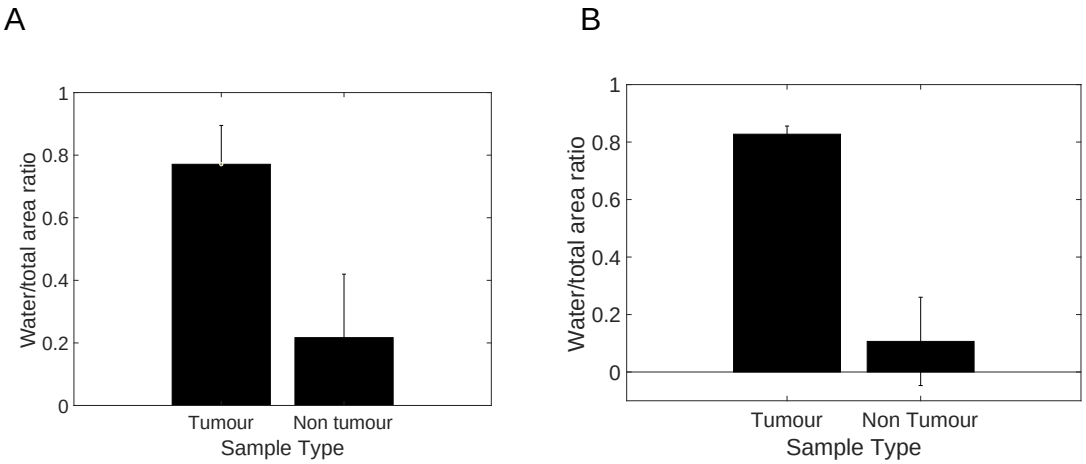


Figure 4-8 Bar charts illustrating the difference in water/total area ratio between tumour (n=96) and non-tumour specimens (n=96)

A Bar graph demonstrating the mean water/total area ratio per specimen of tumour and non-tumour specimens, error bars represent +/- 1 S.D. Statistically significant difference between the mean water/total area ratio of tumour (0.7709; SD +/- 0.1240) versus non-tumour (0.2169; SD +/- 0.2027) calculated by student T-Test ($P=1.6X10^{-42}$)

B Bar graph demonstrating the median water/total area ratio per specimen of tumour and non-tumour specimens, error bars represent +/- interquartile range. Statistically significant difference between the median water/total area ratio of tumour (0.827; IQR – 0.0284) versus non-tumour (0.1065; IQR – 0.1536) calculated by Kruskal Wallis test ($P= 2.3 X 10^{-39}$).

Having established there was a difference in water/total area ratio between tumour and non-tumour specimen spectra, the aim was to investigate if there was a ratio threshold that could predict if a spectrum corresponded to tumour or non-tumour tissue, i.e. if a spectra had a ratio above a certain value it would be classified as coming from tissue containing tumour. This could then be used to differentiate between normal and cancerous tissue and be a foundation of intra-operative margin analysis. Binomial logistic regression was performed with the water/total area ratio with the known classifier of whether the ratio was obtained from tumour or non-tumour tissue. A Receiver Operator Characteristic (ROC) curve (Figure 4-9) was constructed to obtain this ratio threshold value – which was the point on the curve that related to the optimal balance between sensitivity and specificity for diagnosis. The water/total area ratio from all the spectra as individual data points (n=1335) was used to construct ROC curves, and a classification trees prediction model with a two way split, which were then cross-validated with a number of techniques, the results are shown in (Table 4-1). All binomial logistic regression models with cross validation showed good sensitivity

of between 92-94%, specificity of 86-87% with an overall accuracy of 89%, with the classification trees model having a reduced sensitivity leading to a lower overall accuracy.

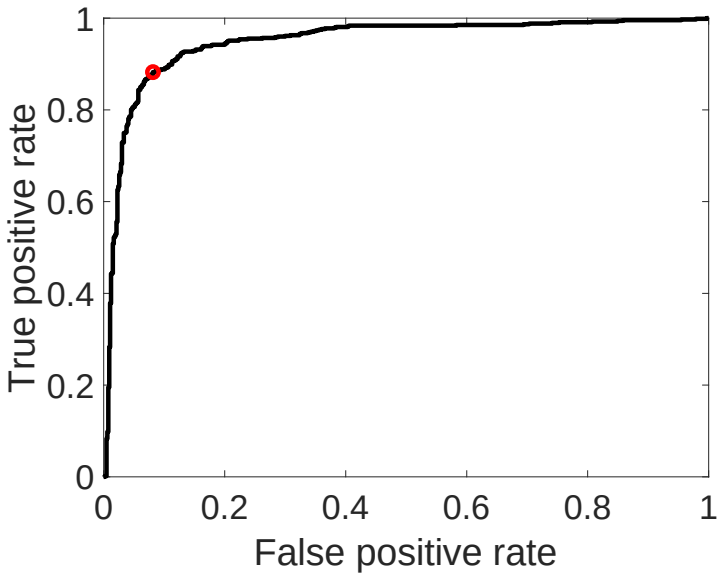


Figure 4-9 ROC curve for classification by binomial logistic regression of water/total area ratio prediction of tumour versus non-tumour of all spectra (n=1335)

Red circle denotes the optimal water/total area ratio threshold for classification = 0.75

Model	Optimal Threshold	AUC	Cross validation technique	Sens.	Spec.	Overall accuracy
Binomial logistic regression	0.75	0.95	n/a	92.7	86.4	89.6
	0.56	0.95	Hold out	91.5	87.4	89.5
	0.77	0.95	K folds 5	93.1	86.6	89.9
	0.76	0.95	K fold 10	92.7	86.4	89.6
	0.76	0.95	LOOCV	92.7	86.6	89.7
Classification Trees	0.68	0.95	LOOCV	87.5	91.7	89.6

Table 4-2 Summary table of results following Model predictions of categorising spectra as tumour versus non-tumour based on water/total area ratio in all spectra (n=1335)

A possible limitation of using every spectrum in the analysis is that multiple spectra from each specimen is used in the construction of the prediction model. This means that if a specimen had a large number of spectra taken from it, or had some outlying ratio values, this may disproportionately influence the model and limit the accuracy of the threshold value. Therefore, the same sequence of analysis was undertaken with a single value for each specimen. The value was the mean of all the water/total area ratio for each specimen. Multiple ROC curves were constructed using binomial logistic regression and a classification tree with a two way split (a representative ROC curve is shown in Figure 4-10) with a number of different cross validation methods (Table 4-3). There was improved accuracy using the mean spectral data compared to the individual data, with an overall accuracy of the diagnostic test of 89-92%.

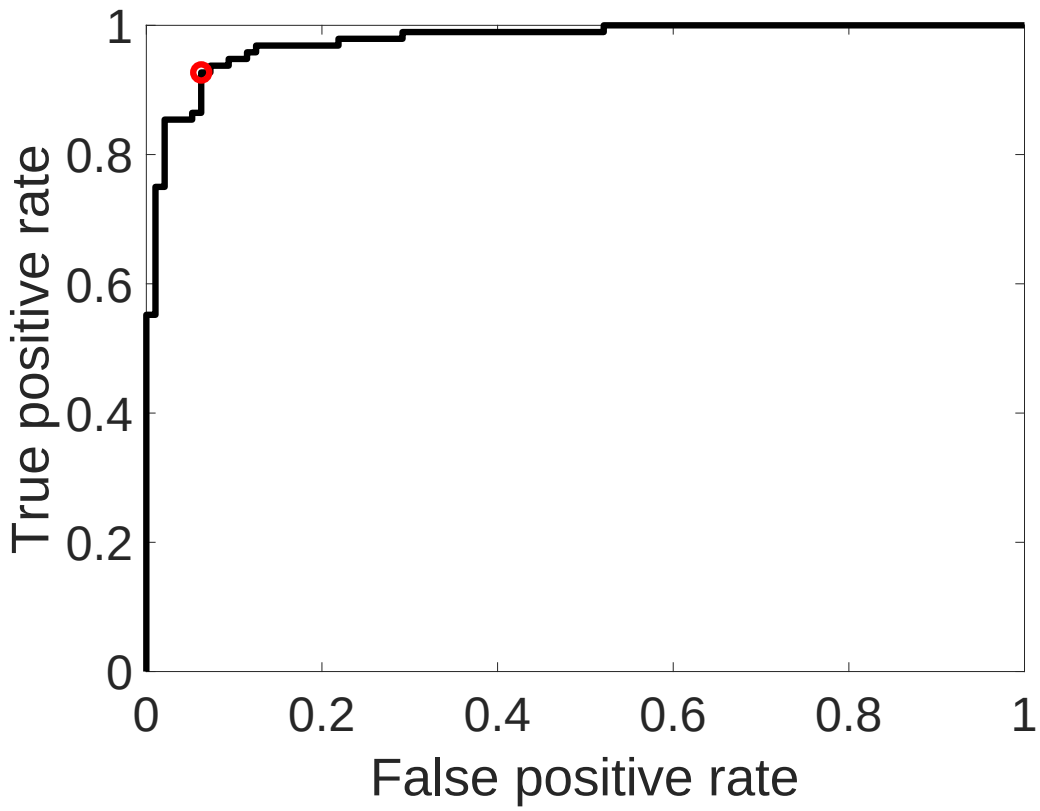


Figure 4-10 ROC curve for classification by binomial logistic regression of water/total area ratio prediction of tumour versus non-tumour of mean spectra (n=192)

Red circle denotes the optimal threshold for classification, which is 0.6264 in this example

Model	Optimal Threshold	AUC	Cross validation technique	Sens.	Spec.	Overall accuracy
Binomial logistic regression	0.6264	0.98	n/a	95.8	88.5	92.2
	0.83	0.95	Hold out	92.9	86.2	89.5
	0.6587	0.97	K folds 5	94.8	89.6	92.2
	0.6333	0.97	K fold 10	94.8	88.5	91.7
	0.6524	0.97	LOOCV	94.8	88.5	91.7
Classification Trees	0.563	0.93	LOOCV	92.7	90.6	91.7

Table 4-3 Summary table of results following Model predictions of categorising spectra as tumour versus non-tumour based on water/total area ratio in mean spectra (n=192)

Voting thresholds

A clinically relevant method of processing spectral data would be to analyse all the spectra obtained from the resected edge of a specimen and determine whether there was tumour present at the resected edge. This would provide a simple dichotomous output to inform the surgeon that the margin was 'positive' and that the tumour bed of the resected edge required resecting. A way of providing this output is to determine that if the majority of the resected edge had a 'tumour' water/total area ratio (as defined by the threshold ratio set by binomial logistic regression e.g. 0.75), then the entire resected edge would be classified as tumour. The water/total area ratio data was analysed in an algorithm that stipulated that if the majority of the specimen measurements had a 'tumour' water/total area ratio, the whole specimen was classified as tumour (Voting threshold 1).

A second algorithm was used on the same data that stipulated if any of the specimen demonstrated a single 'tumour' water/total area ratio, the whole specimen would be determined as tumour (Voting threshold 2).

The results of these two algorithms showed that voting threshold 1 gave a high degree of specificity (95.8%), but a lower sensitivity of 87.5% with a good overall accuracy of 91.7%. Voting threshold 2 had a higher degree of sensitivity of 95.98%, but a low specificity (88.0%) of and an overall accuracy was 88.0 % (Figure 4-11).

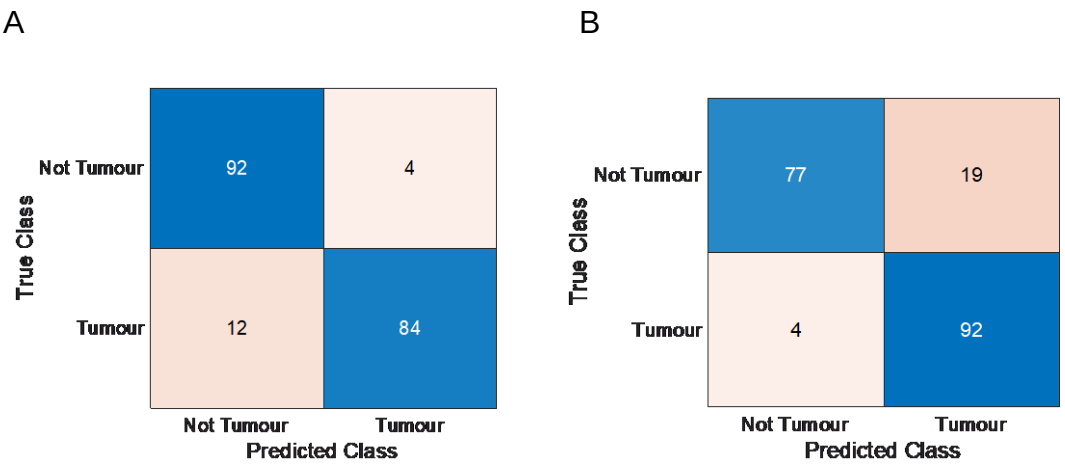


Figure 4-11 Confusion matrices of specimen diagnosis after voting threshold analysis of water/total area ratio using the optimal threshold of water/total area ratio of 0.75

A. Voting threshold 1 which achieved overall accuracy – 91.7%, sensitivity – 87.5%, specificity – 95.8% B. Voting threshold 2 which achieved overall accuracy – 88.0%, sensitivity – 95.8%, Specificity – 80.2%.

The threshold used to classify as a 'tumour' signal in these voting thresholds could be manipulated to tend towards being more sensitive or more specific based on what is necessary when analysing an entire margin. For example, if a water/total area ratio threshold of 0.4 is used in voting threshold 1, the algorithm has a 90.1% accuracy, with 97.9% sensitivity and 82.3% specificity, and if a water/total area threshold of 0.8 is used in voting threshold 1, the algorithm has a 89.1% accuracy, with 79.2% sensitivity, and a 99.0% specificity Table 4-4.

Ratio Threshold	Sensitivity	Specificity	Accuracy
0.1	99.0	44.8	71.9
0.2	99.0	70.8	84.9
0.3	97.9	79.2	88.5
0.4	97.9	82.3	90.1
0.5	95.8	86.5	91.1
0.6	93.8	90.6	92.2
0.7	93.8	93.8	93.8
0.8	79.2	99.0	89.1
0.9	1	99.0	<50

Table 4-4 Table demonstrating the change in diagnostic accuracy with a change in water/total area ratio threshold (Ratio Threshold) to classify the majority of a specimen using voting threshold 1.

Using the intra-patient difference in water/total area ratio

The methods illustrated above, to provide differentiation between tumour versus non-tumour, have used a water/total area ratio threshold based on a value obtained from all spectra from all individuals. The physiological 'normal' in terms of water content for each patient is known to be different (and is explored in a later section of this chapter for these patients). Therefore, having the same cut-off value for all specimens could lead to some diagnostic inaccuracy. A measure of the change in the water/ total area ratio within a patient between tumour and non-tumour could be sufficient to differentiate areas of tumour versus non-tumour. Logistically this could be achieved by the surgeon 'calibrating' the instrument on some obviously normal tissue, well away from the tumour site, prior to examining the specimen for IMA, and areas of difference compared to the 'calibration' measurement flagged as abnormal and requiring resection.

A simple and effective measure of this change is to calculate the difference in the water/total area ratio between the area that is being measured and the 'calibration' measurement from the normal tissue. The diagnostic ability of the water/total area ratio difference to diagnose breast cancer was investigated using the water/total area ratio from all spectra (n=1335). For each patient, the mean

water/total area ratio was calculated from the spectra obtained from the non-tumour specimen – this was the 'calibration' measurement. Then each individual spectra (from tumour and non-tumour specimens) for each patient had the water/total area ratio calculated, and the difference from that reading taken from the 'calibration' measurement as per Equation 4-2:

$$W/TAR\ difference = W/TAR\ of\ current\ measurement - 'Calibration'\ W/TAR$$

Equation 4-2 – Calculation of the water/total area ratio difference

W/TAR = Water/Total area ratio

'Calibration' W/TAR = the mean of all the water/total area ratio for that patient's non-tumour specimen readings

This provides a unique measure for each reading that represents the difference from the mean non-tumour specimen. It could be presumed that non-tumour measurements would be approximately 0 (as they produce the 'calibration' ratio), and tumour measurements would be substantially greater than 0. The cut off W/TAR difference of 0.1 was chosen as it is close to 0 and allows for some intra-specimen variation of non-tumour specimens W/TAR difference, but would likely give a model with high sensitivity. Therefore, for all spectra, the W/TAR difference was calculated, and if the W/TAR difference was >0.1, it was classified as a tumour, and if <0.1, it was classified as non-tumour, and a confusion matrix produced (Figure 4-12). It shows that using this method, the overall accuracy is 89.5%, with a high sensitivity of 93.6% but low specificity of 85.4%.

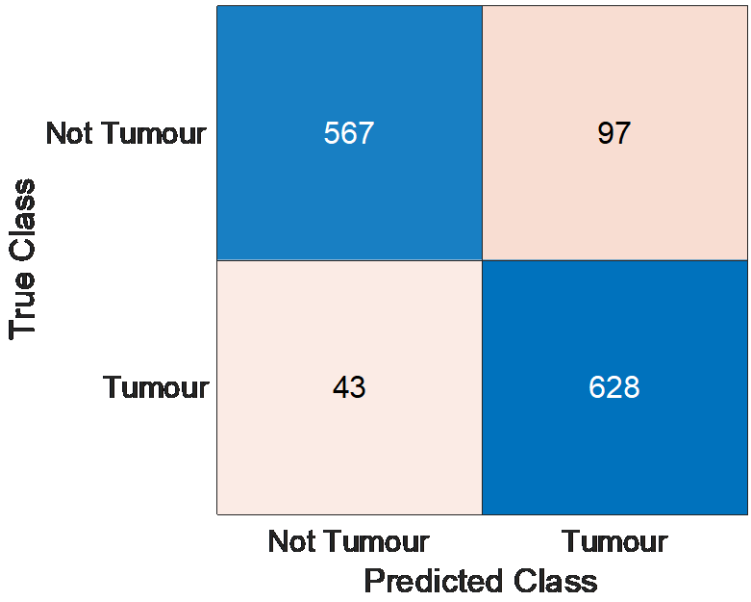
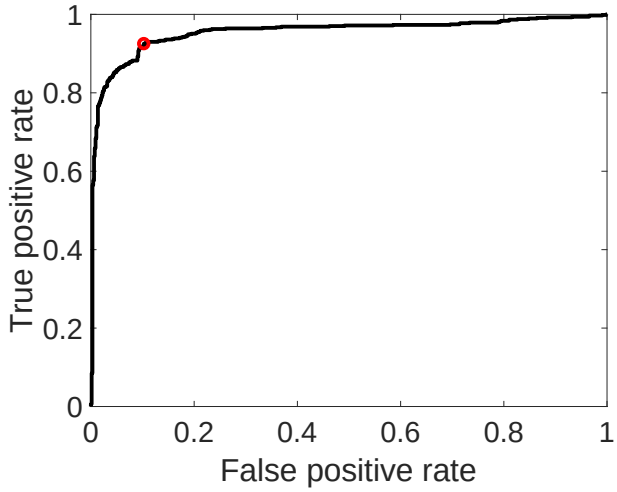


Figure 4-12 Confusion matrix showing the diagnostic results of the W/TAR difference cut off= 0.1 for classifying tumour versus non-tumour spectra

An improved diagnostic accuracy may be possible, and so binomial logistic regression was used using the W/TAR difference calculated for each spectrum, and the classifier of whether it came from a tumour or non-tumour specimen used to train the model. The resultant ROC curve is shown, which produced an AUC of 0.95, and predicted the optimal threshold (cut off for the W/TAR difference) was 0.3275 (if the W/TAR difference >0.3275, the measurement was classified as a tumour reading, and if <0.3275, it was classified as a non-tumour reading) (Figure 4-13). Using a k-folds 5 cross validation method, the model achieved an accuracy of 90.1%, a sensitivity of 87.0% and specificity of 93.2%. The results of using the 'optimal threshold' of 0.3275 (high specificity) compared to the previously chosen threshold of 0.1 (high sensitivity) demonstrate that the threshold can be manipulated to provide the desired test characteristics, similar to using the more simple water/total area ratio calculation.

A



B

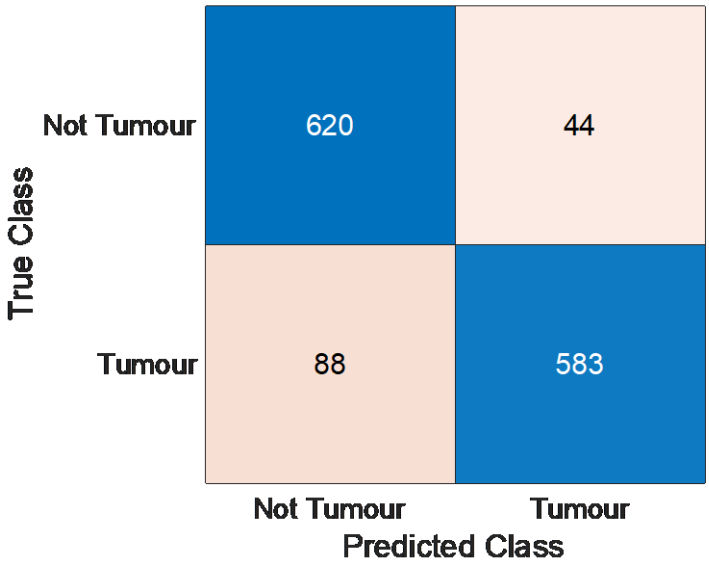


Figure 4-13 Results of binomial logistic regression model with k folds 5 cross validation for predicting tumour versus non-tumour specimen using the W/TAR difference

A. ROC curve for classification by binomial logistic regression of W/TAR difference prediction of tumour versus non-tumour of all spectra (n=1335). Red circle denotes the optimal threshold for classification, which is 0. 3275. B. Confusion matrix of prediction results after k folds 5 cross validation

Another way of performing this analysis to perform IMA would be to take a number of readings over the area of the margin and calculate the mean water/total area ratio for the entire margin surface prior to subtracting the mean 'calibration' reading and providing a single output for that area. The mean water/total area ratio from each patient's 'non-tumour' specimen was subtracted from the matched patient's 'tumour' specimen (as representative of values that might be obtained along the margin), to give the water/total area ratio difference for that patient's 'tumour' specimen (Figure 4-14). This shows that for the mean readings, the water/total area ratio difference was greater than 0 in all but 1

patient, suggesting that if a threshold water/total area ratio difference of 0 was set, 95 of 96 of patients would be correctly classified. This cannot be assessed in more detail for diagnostic accuracy, as the mean non-tumour tissue value is by definition 0, and so prediction models and cross validation cannot be performed.

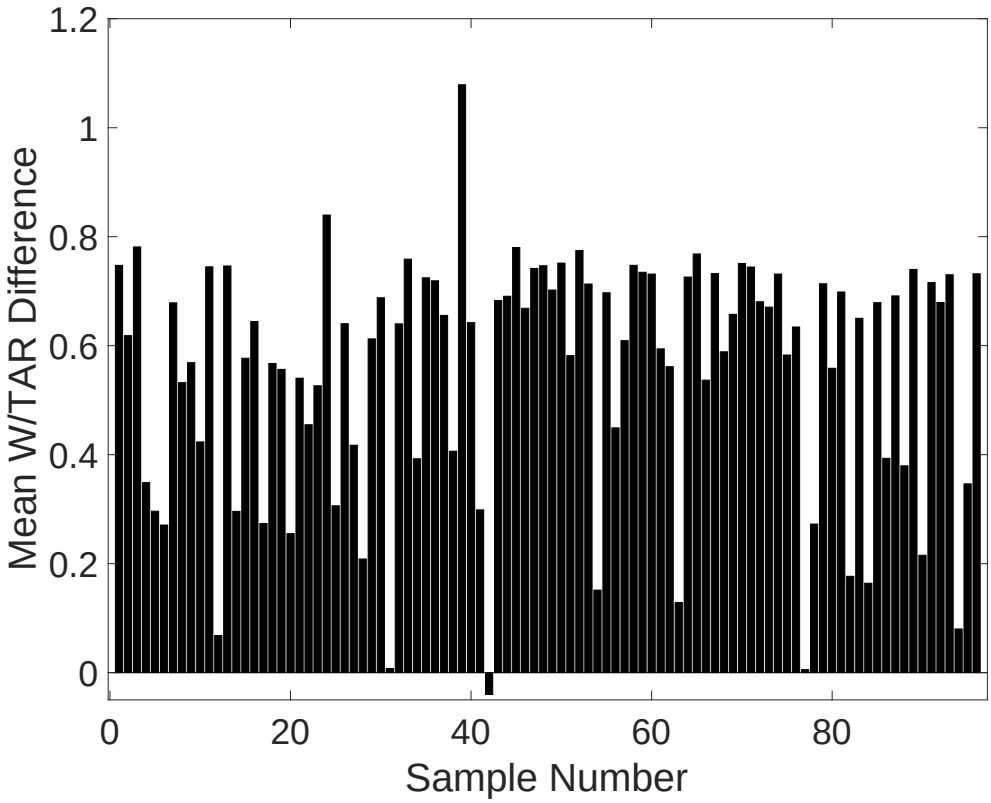


Figure 4-14 Bar chart of sample number (n=96) and the mean W/TAR Difference for the tumour specimen for that sample

Note – only one sample has a negative value, suggesting that if a cut off of W/TAR Difference of 0 was used, all but one sample would be classified as tumour.

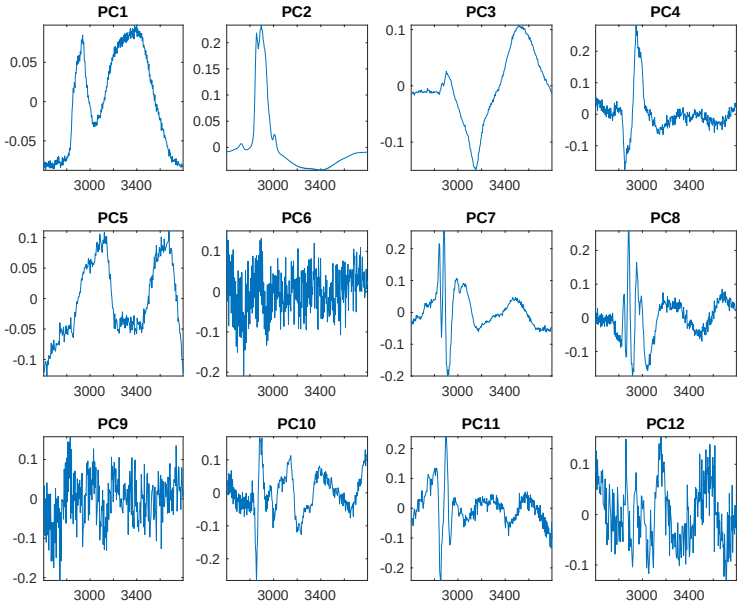
4.3.2.3 Spectral analysis to diagnose breast cancer

HWN region

PCA analysis of pre-processed HWN spectral data was performed, with the results demonstrated in Figure 4-15. Two of the three principal components that achieved statistical significance between tumour and non-tumour tissue (PC1 and PC2) had incredibly small P values (recorded as 0, as they were so low as to be rounded to 0). These two PCs were confirmed as accounting for the majority of the classification, as PC1 accounted for 66% of the variance, and PC2 33% of the variance, with the remaining 10 PC's accounting for <1% of the variance. Plotting the three most significant principal components (as defined by the smallest P value) gave excellent visual separation between tumour and non-tumour spectra.

PC1 and PC2 were correlated with the spectral assignation from previous work in this thesis and were shown to correspond to protein and water peaks (PC1), and the lipid peak (PC2), which also correlated with the spectral assignation from the literature [295, 304] (Figure 4-16). This suggests that it is both of these spectral features that allow separation of spectra according to the tumour or non-tumour classification, and are in keeping with the findings from the water/total area ratio calculations, that tumour (protein and water rich) can be differentiated from non-tumour (lipid rich, with low water concentration) using the HWN region.

A



B

PC 1	PC 2	PC 3	PC 4
0	0	0.97	0.86
PC 5	**PC 6**	**PC 7**	**PC 8**
0.91	0.39	0.50	0.0026
PC 9	**PC 10**	**PC 11**	**PC 12**
0.24	0.52	0.11	0.99

C

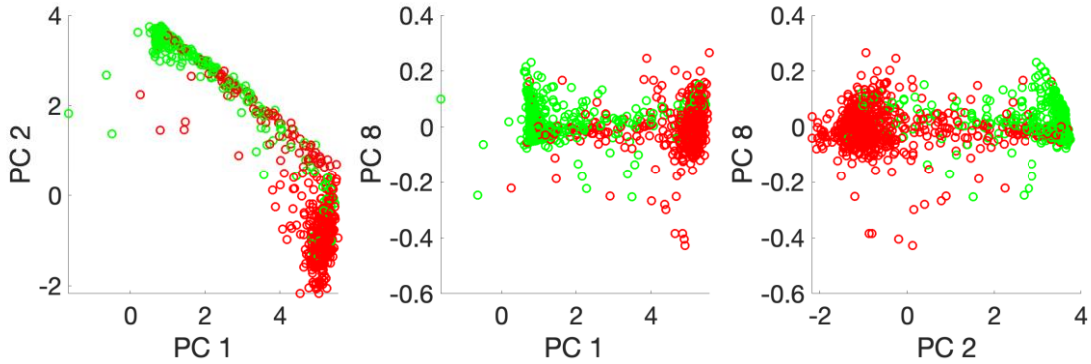


Figure 4-15 Principal Component Analysis of all HWN spectral data (n=1335)

A. Principal Components (12 in total) identified in analysis B. Table of principal components and corresponding P Values C. Scatter graphs plotting the 3 significant Principal Components (PC1, PC2, PC8), with tumour scores plotted in red, non-tumour scores in green

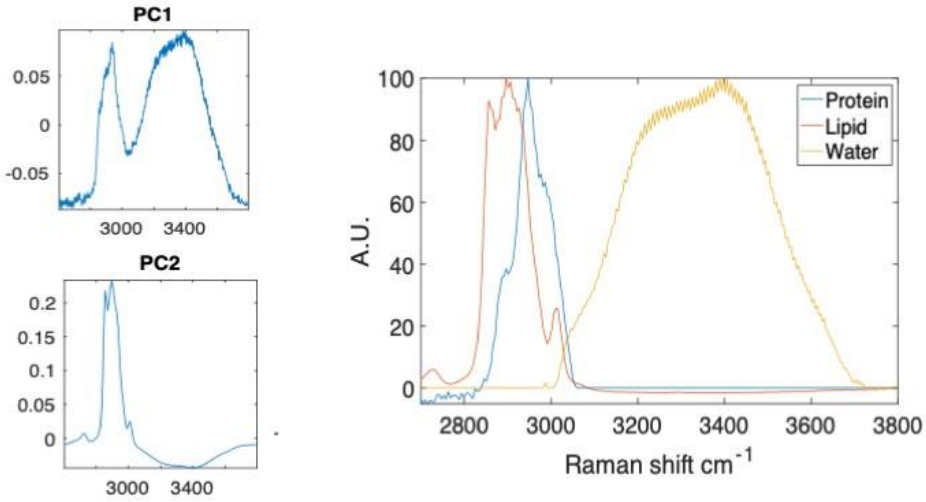
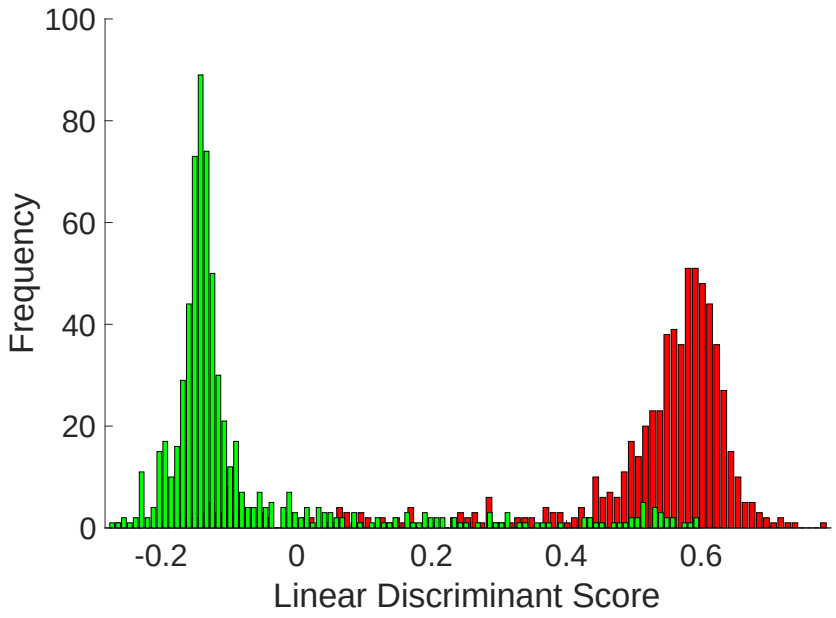


Figure 4-16 Principal components accounting for the majority of spectral variance between tumour versus non-tumour tissue, and spectral assignation

Graphs on Left are the spectrum of Principal Components 1 and 2, and on the Right are representative Raman spectrum of bovine gelatine (protein), soya bean oil (lipid) and distilled water (water). It can be seen that PC1 relates to protein and water signal, and PC2 relates to lipid signal, suggesting it is these components that differentiate between tumour and non-tumour tissue.

The PCA scores were fed into a LDA which achieved good separation between spectra and a training performance of 90.5%, and subsequently a LOOCV performed, the results of which are in Figure 4-17. After LOOCV the technique obtained sensitivity 90.2%, Specificity 90.5% with an overall accuracy of 90.3% (Figure 4-17).

A



B

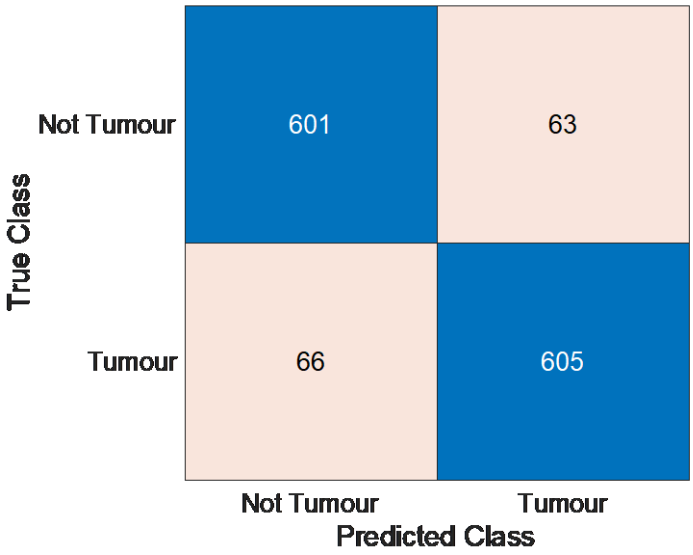
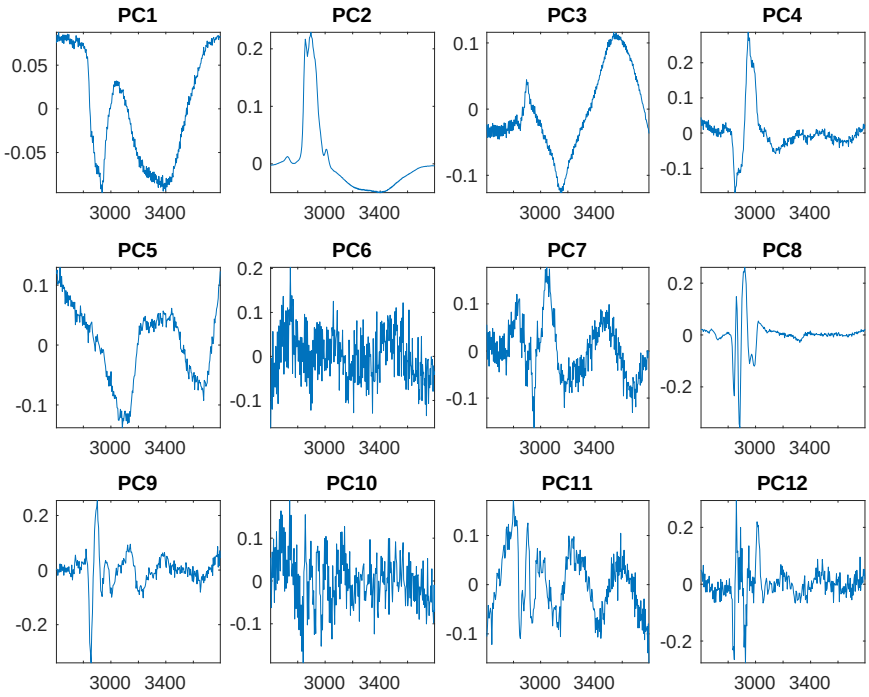


Figure 4-17 Results of all HWN spectra (n=1335) to compare spectra from tumour versus non- tumour specimens with PCA fed LDA analysis

A. Histogram of LDA scores versus frequency, with tumour scores plotted in red, non-tumour scores in green. B. Confusion matrix of prediction after LOOCV

The mean spectrum for each specimen was calculated, providing a single spectrum for each specimen (n=192). PCA was performed which revealed similar results – PC 1 and PC2, relating to the protein and lipid peak had significant P values, with PC1 accounting for 67.4% of variance, and PC2 accounting for 32.1%, with the remaining scores combined accounting for only 0.3% of the variance (Figure 4-18).

A



B

PC 1	PC 2	PC 3	PC 4
$5.0X10^{-64}$	$4.0X10^{-65}$	0.97	0.79
PC 5	**PC 6**	**PC 7**	**PC 8**
0.81	0.80	0.56	0.53
PC 9	**PC 10**	**PC 11**	**PC 12**
0.72	0.97	0.66	0.82

C

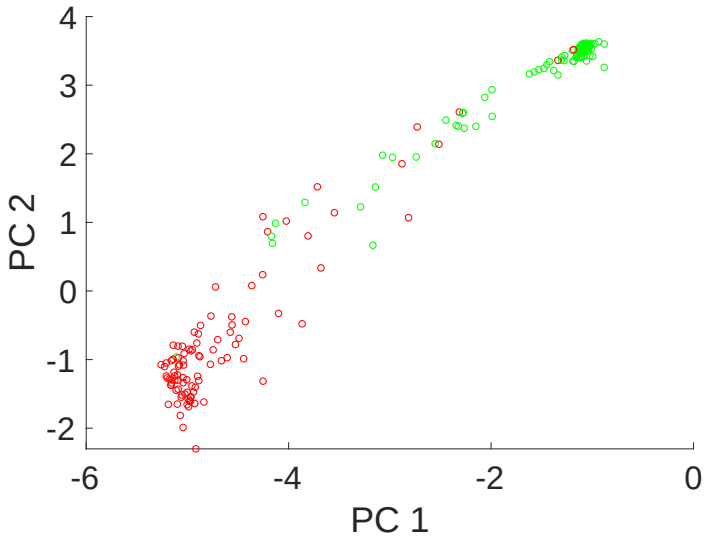


Figure 4-18 Principal Component Analysis of mean HWN spectral data (n=192)

A. Principal Components (12 in total) identified in analysis B. Table of principal components and corresponding P Values C. Graph plotting the 2 significant Principal Components (PC1 and PC2), with tumour scores plotted in red, non-tumour scores in green.

LDA analysis gave excellent separation with a training performance of 93.8% accuracy, and LOOCV gave a sensitivity of 93.8%, specificity 92.7% and overall accuracy 93.2% (Figure 4-19).

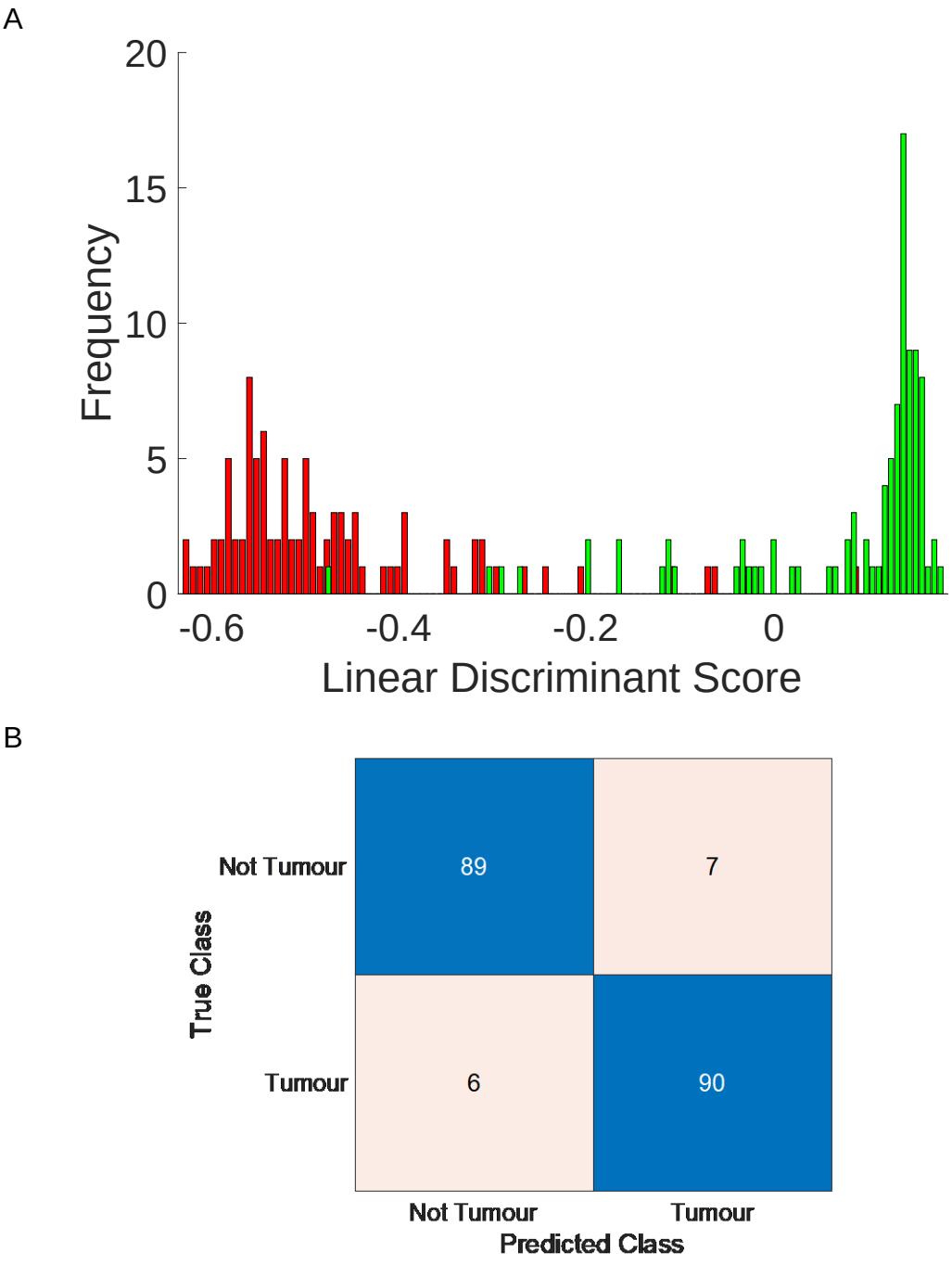


Figure 4-19 Results from spectral analysis of the mean HWN spectra for each specimen (n=192)
A. PCA fed LDA histogram of scores versus frequency using 12 principal components with tumour scores plotted in red, non-tumour scores in green. B. Confusion matrix prediction of LDA LOOCV.

Fingerprint region

Spectra from the FP region underwent pre-processing, the mean spectra from tumour and non-tumour specimens is shown in Figure 4-20. There are evident spectral differences in the FP region between spectra obtained from tumour specimens and non-tumour specimens.

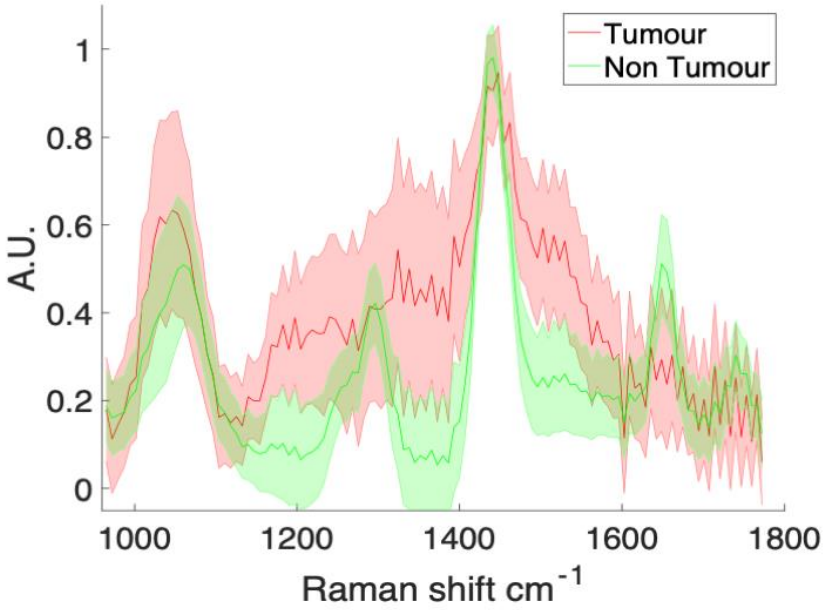
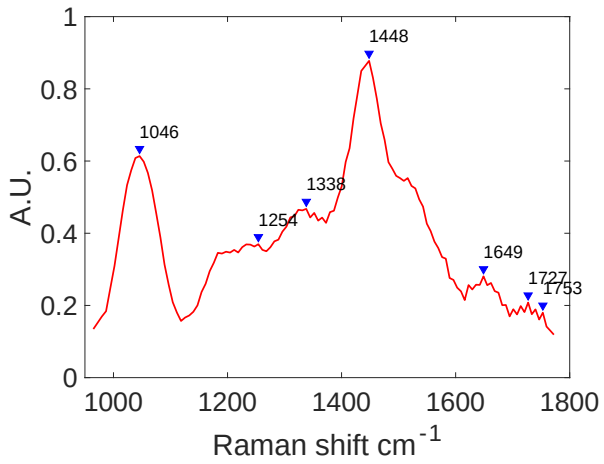


Figure 4-20 Mean Fingerprint spectra of all specimens (n=1335) according to diagnosis demonstrating differences between tumour and non-tumour specimen spectra

Data is mean of spectra from all tumour specimen spectra (n=672) and all non-tumour specimen spectra (n=663), fluorescent baseline subtraction with a 6th order polynomial and min/max normalisation to avoid distortion in peak heights. Shading either side of solid line in same colour is +/- 1 S.D. Red spectra=measurement from tumour specimen, Green spectra = measurement from non-tumour specimen

To understand the spectral differences between tumour and non-tumour specimens, the two groups of mean spectra were separated, peaks identified, and spectral peak assignation performed based on values from the literature (Figure 4-21). This demonstrated that the differences between them primarily related to differences in protein and lipid signals. This suggests that the FP region may differentiate between tumour and non-tumour based on the differences between protein and lipid tissue, which is the same basis for differentiation in the HWN region.

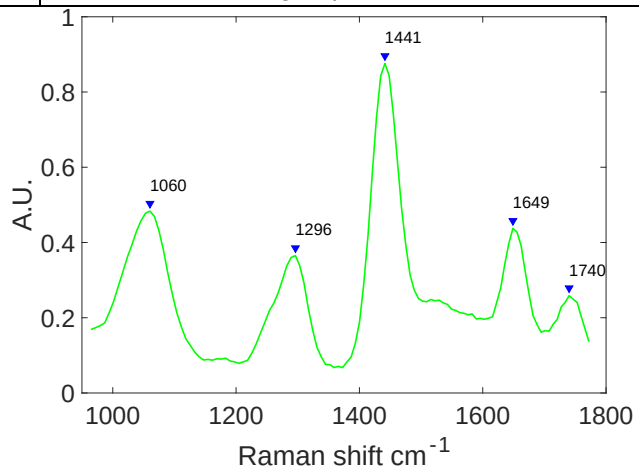
A



B

Spectral Peak	Biochemical assignation
1046 (1034-1053)	C-O stretching, C-N stretching (protein) Possibly collagen
1254 (1200-1300)	Amide III (proteins)
1338 (1335-1345)	Amide III; CH_3CH_2 wagging mode of collagen
1448	CH_2 & CH_3 deformation; collagen
1649	Amide I (C=C)
1727-1753	C=O stretch, ester group

C



D

Spectral Peak	Biochemical assignation
1060	Skeletal C-C stretch DNA/Lipid
1296	CH_2 deformation
1441	CH_2 scissoring & CH_3 bending in Lipids
1649	Lipid (C=C stretch)
1740	C=O stretch Lipid, ester group

Figure 4-21 Spectral assignation of Fingerprint spectra in breast tissue

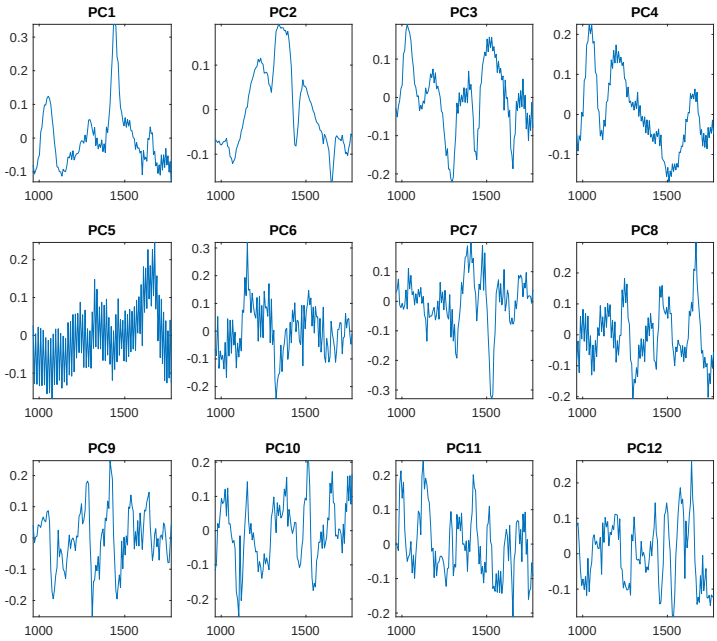
A. Mean of all spectra taken from tumour tissue, with peaks identified with Raman shift number. B. Spectral assignation table for tumour tissue peaks. C. Mean of all spectra taken from non-tumour tissue. D. Spectal assignation table for non-tumour tissue peaks.

The mean Raman spectra have been smoothed with savitzky-golay span of 5, 1st order polynomial and the main spectral peaks are identified with corresponding Raman shift value. Biochemical peak assignation is according to the literature [103] [356].

PCA analysis was performed and over half of the principal component scores reached statistical significance in classifying tumour versus non-tumour scores (7 of 12), which is more than that of the HWN region (Figure 4-22). The first 3 PCs were confirmed as being responsible for the majority of the difference between the groups. Variance analysis of the principal components demonstrated that PC 1 accounted for 58.5% of the variance, and PC 2 25.4 % of the variance, with PC 3 accounting for 10.59% of the variance and the remaining 9 PC's accounting for 5.5% of the variance.

It can be seen that in plotting the three most significant principal components (as defined by the smallest P value), it gave good visual separation between tumour and non-tumour spectra.

A



B

PC 1	PC 2	PC 3	PC 4
$4.08X10^{-27}$	$4.89X10^{-115}$	$2.89X10^{-116}$	0.99
PC 5	**PC 6**	**PC 7**	**PC 8**
$1.93X10^{-6}$	$3.56X10^{-4}$	$6.30X10^{-6}$	0.02
PC 9	**PC 10**	**PC 11**	**PC 12**
0.032	0.06	$6.56X10^{-4}$	0.66

C

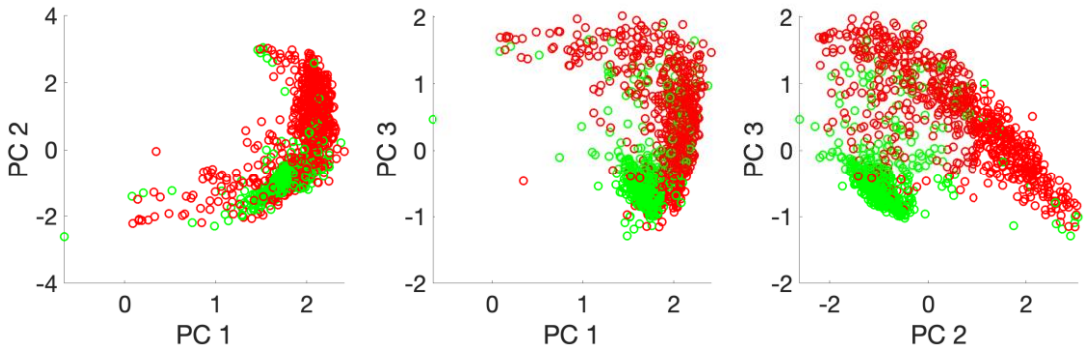


Figure 4-22 Principal Component Analysis of Fingerprint region spectra of all spectra (n=1335)

A. Principal Components (12) identified in analysis B. Table of principal components and corresponding P Values C. Scatter graphs plotting the 3 most significant Principal Components (PC1,PC2,PC3) accounting for >95% of variance, with tumour data plotted in red, non- tumour data in green

The spectral peaks, with the Raman shift values, were identified for the spectral features of PC1 and PC2 as these accounted for 85% of variance. The peaks of PC1 relate to the peaks seen for tumour tissue, and the peaks of PC2 relate to the peaks seen for non-tumour tissue (Figure 4-23).

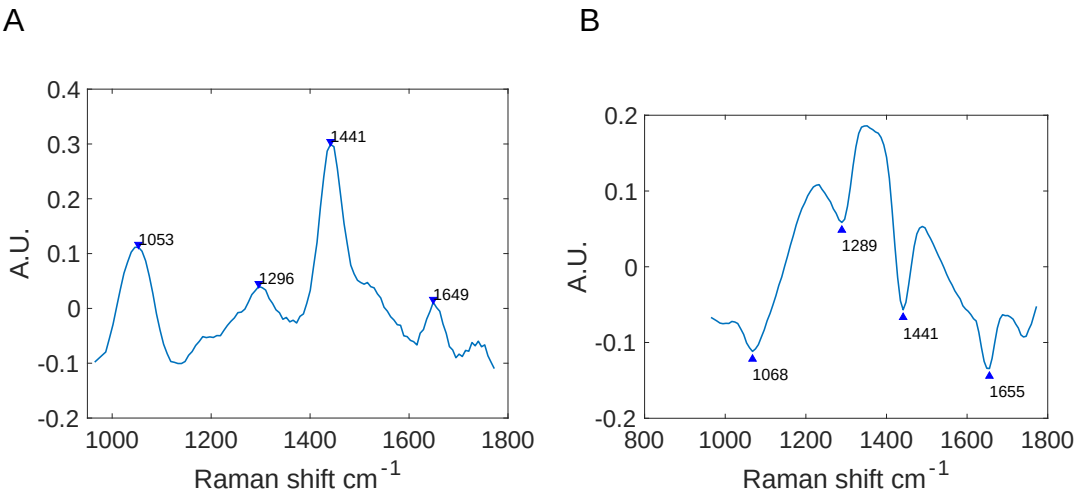


Figure 4-23 Spectral assignation in the two most significant principal components in the Fingerprint region in differentiating tumour from non-tumour tissue

A. Principal Component 1; these peaks correspond to those identified as corresponding to tumour tissue (protein rich) B. Principal Component 2; these peaks correspond to those identified as corresponding to non-tumour tissue (lipid rich)

Peak assignments of all of the principal component Raman peak scores that reached statistical significance, along with a tentative assignment to the molecule type is in Table 4-5. This demonstrates that most of the differentiation between tissue types is due to protein and lipid differences, with some minor variance due to DNA/RNA. This suggests, along with the mean spectra data analysis above, that the differentiation of tumour versus non-tumour tissue in the fingerprint region is predominantly based on the difference in tumour (protein rich) and non-tumour (lipid rich). This is similar to the HWN region, however, in the fingerprint region there is no measure of water content differences between tissue types.

Principal component	Peaks (Raman shift cm^{-1})	Biochemical assignation	Overall tentative assignment
1	1053	C-O stretching, C-N stretching (protein)	Protein
	1296	Amide III (proteins)	
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1649	Amide I (C=C)	
2	1068	Skeletal C-C stretch DNA/Lipid	Lipid
	1289	CH_2 deformation	
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1655	Lipid (C=C stretch)	
3	1296	CH_2 deformation	Lipid
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1655	Lipid (C=C stretch)	
5	1675	Amide I (Beta sheet)	Protein
6	1154	C-C and C-N stretch protein	Protein
7	1324	CH_3 and CH_2 wagging of purine bases	DNA
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1529	Carotenoids or Cytosine	
11	994	C-O Ribose	DNA/RNA
	1126	C-C stretch protein or carbohydrates	
	1421	CH deformation DNA/RNA/Deoxyribose	
	1516	Cytosine or carotenoids	

Table 4-5 Peak assignment of significant Principal component scores in fingerprint spectral analysis
Biochemical peak assignation is according to the literature [103] [356]

LDA analysis using the scores from the PCA gave a training performance of 90.6%. LOOCV analysis of the model gave an overall accuracy of 90.3%, sensitivity 90.2% and specificity 90.5% (Figure 4-24).

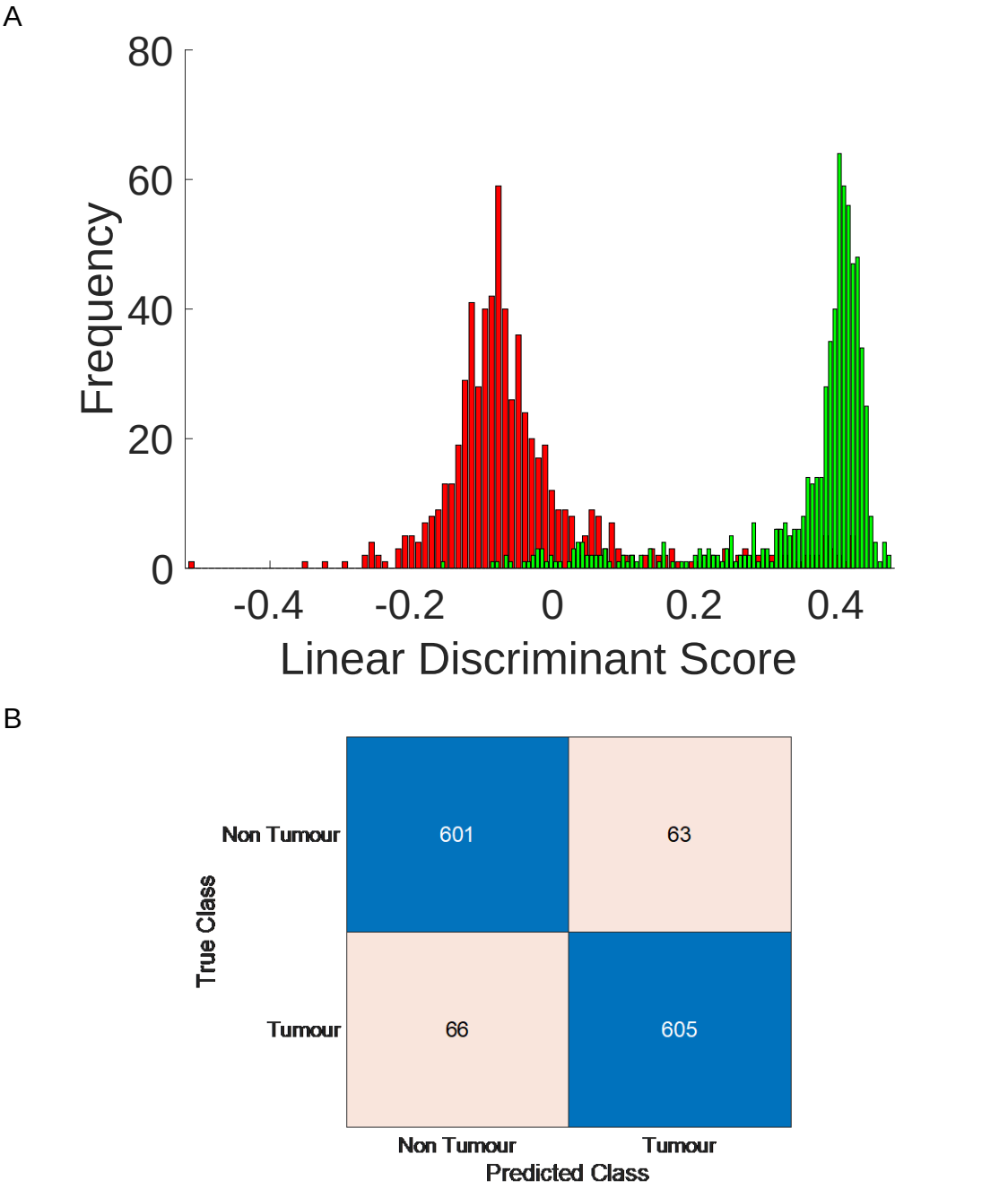
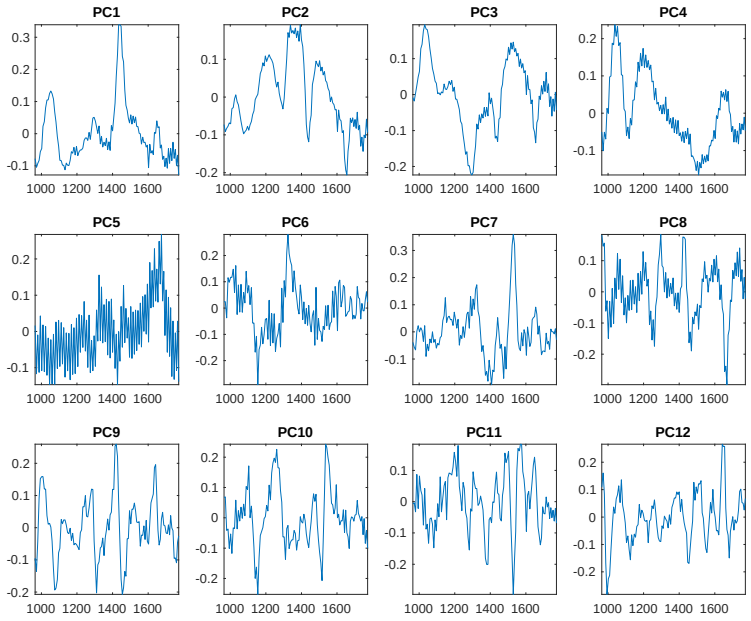


Figure 4-24 PCA fed LDA analysis of all Fingerprint region spectral data (n=1335) to differentiate between tumour and non-tumour breast spectra

A. Histogram of LDA scores versus frequency, with tumour data plotted in red, non-tumour data in green B. Confusion matrix of prediction after LOOCV

The mean spectrum for each specimen was calculated and underwent PCA analysis. This gave visually similar principal components for the first 12 PC's to the individual spectra data PC 1 accounted for 68.6% of variance, PC 2- 21.0% so PC1 and PC2 accounted for >90% of the variance, with PC 3 accounting for 7.26% and PC 4-12 – 3.11% (Figure 4-25).

A



B

PC 1	PC 2	PC 3	PC 4
$7.90X10^{-238}$	$1.97X10^{-26}$	0.089	$1.32X10^{-5}$
PC 5	**PC 6**	**PC 7**	**PC 8**
0.049	0.0011	0.0013	$4.12X10^{-6}$
PC 9	**PC 10**	**PC 11**	**PC 12**
0.0091	0.0258	$8.9X10^{-4}$	0.30

C

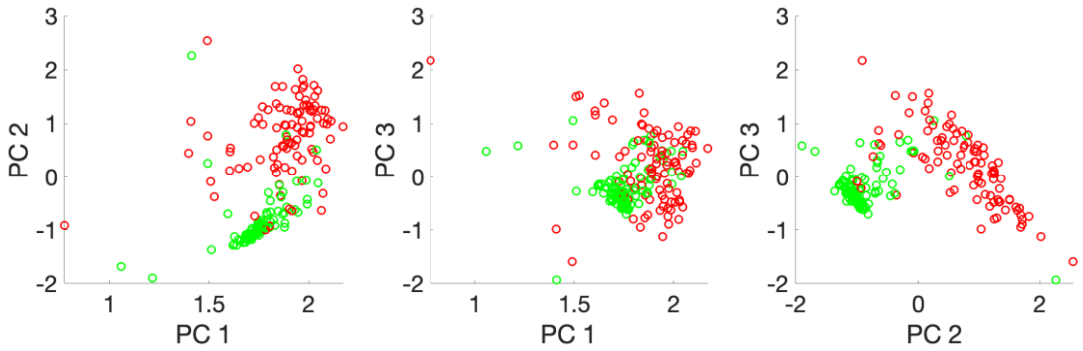


Figure 4-25 Principal Component Analysis of Fingerprint region spectra of mean spectra (n=192)

A. Principal Components (12) identified in analysis B. Table of principal component scores and corresponding P Values C. Scatter graphs plotting the first 3 Principal Components (PC1,PC2,PC3) accounting for >95% of variance, with tumour data plotted in red, non-tumour data in green

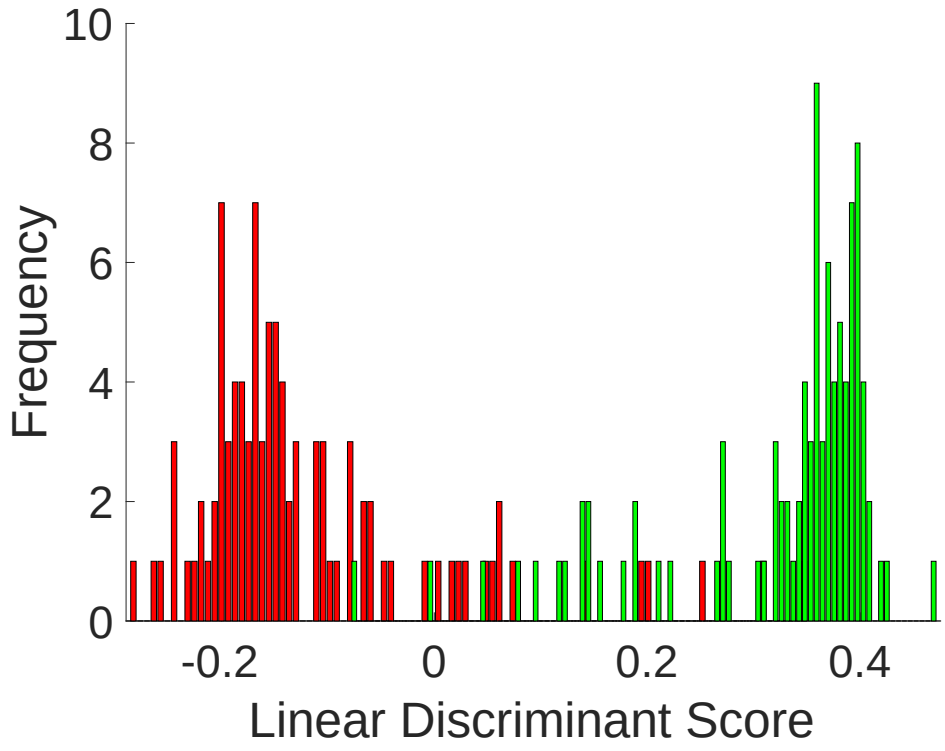
Although the principal components were very similar to the individual spectral data, not all the significant components were the same ones between individual and mean data. Table 4-6 highlights the significant principal components in the mean data, with tentative spectral assignment. It shows that, similar to the individual data, PC 1 and PC2 (that accounted for the majority of the variance) corresponded to protein and lipid peaks, with some other PC's contributing DNA/RNA peaks.

Principal component	Peaks	Biochemical assignation	Overall tentative assignment
1	1053	C-O stretching, C-N stretching (protein)	Protein
	1296	Amide III (proteins)	
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1649	Amide I (C=C)	
2	1068	Skeletal C-C stretch DNA/Lipid	Lipid
	1289	CH_2 deformation	
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1655	Lipid (C=C stretch)	
4	1038	Collagen	Protein
	1197	Amide III	
	1649	Amide I (C=C)	
6	1154	C-C and C-N stretch protein	Protein
7	1324	CH_3 and CH_2 wagging of purine bases	DNA
	1441	CH_2 scissoring & CH_3 bending in Lipids	
	1529	Carotenoids or Cytosine	
8	1261	Amide III	Protein
	1469	C=N stretch	
	1669	C=O stretch; protein band	
9	1075	C-C stretch lipid	Lipid or DNA
	1310	CH stretch lipids	
	1455	Deoxyribose	
11	994	C-O Ribose	DNA/RNA
	1126	C-C stretch protein or carbohydrates	
	1421	CH deformation DNA/RNA/Deoxyribose	
	1516	Cytosine or carotenoids	

Table 4-6 Peak assignment of significant Principal component scores in fingerprint mean spectral analysis Biochemical peak assignation is according to the literature [103] [356]

The scores of the 12 components were fed into LDA analysis which gave a training performance of 94.3% (Figure 4-26). LOOCV was performed which gave 92.7% overall accuracy, sensitivity 91.7% and specificity 93.8%.

A



B

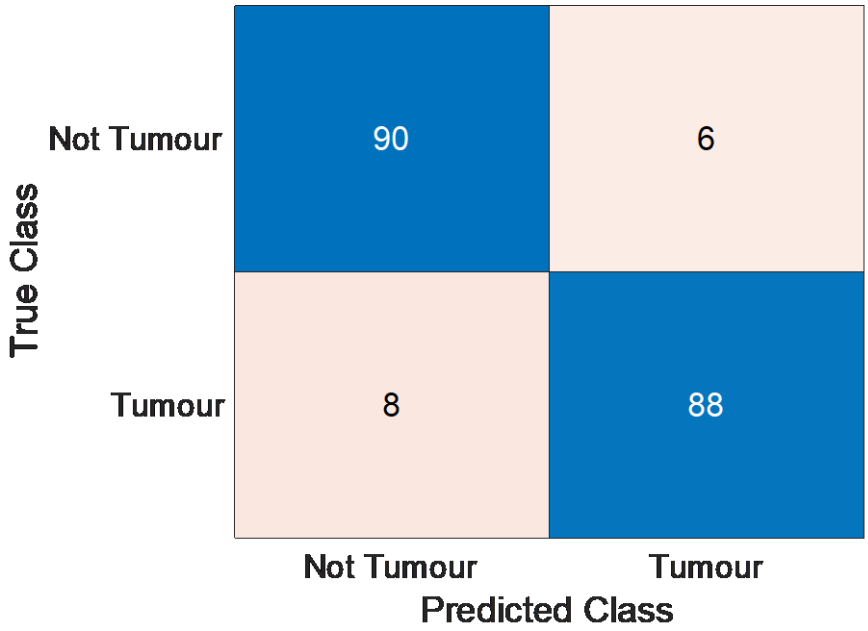


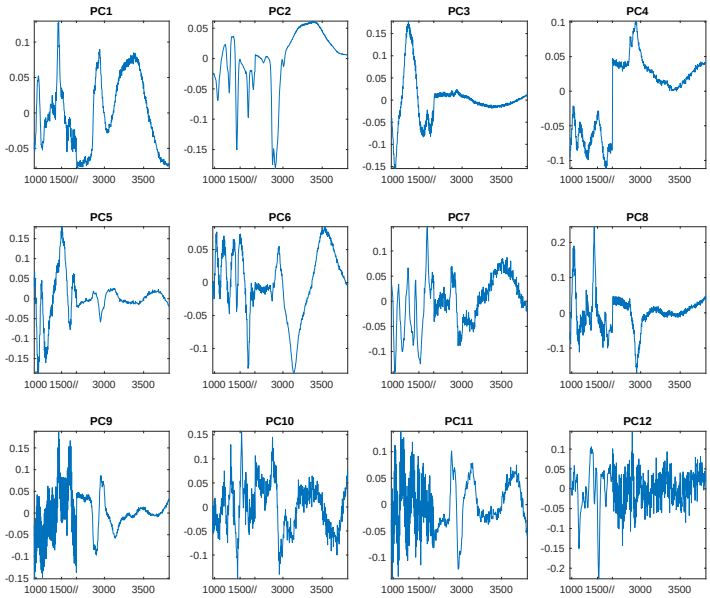
Figure 4-26 PCA fed LDA analysis of mean Fingerprint spectra (n=192) to differentiate tumour and non- tumour specimens

A Histogram of PCA fed LDA scores, with tumour data plotted in red, non-tumour data in green. B. confusion matrix results from LOOCV analysis to predict tumour versus non-tumour

Concatenated spectral data

The spectra from the HWN region and FP region were stitched together and analysed simultaneously. Principal component analysis found 3 components (PC1, PC2 and PC12 with significant P values <0.01 (Figure 4-27). Two of the three principal components (PC 1 and PC2) visually relate to the HWN region, with the protein/lipid peak and water peak being the spectral features responsible for the highly significant scores, as PC1 accounted for 59.8% of the variance, and PC2 33.0% of the variance. PC3 related to the fingerprint region amide III peak at 1338 cm^{-1} associated with protein accounting for 3.8% of the variance with the remaining 9 PCs accounting for the remaining 3.4%.

A



B

PC 1	PC 2	PC 3	PC 4
$1.27X10^{-318}$	0	0.01	0.16
PC 5	**PC 6**	**PC 7**	**PC 8**
0.15	0.75	0.20	0.45
PC 9	**PC 10**	**PC 11**	**PC 12**
0.72	0.01	0.058	$4.0X10^{-4}$

C

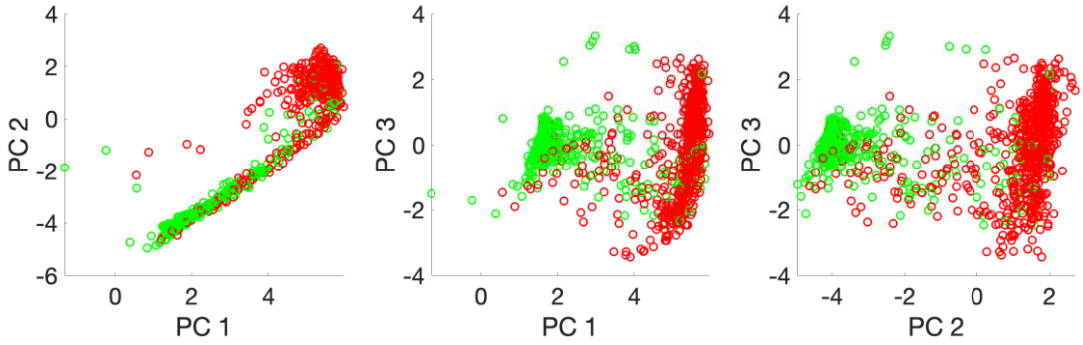


Figure 4-27 Principal component analysis of all concatenated spectral data (n=1335)

A 12 Principal components that were identified and underwent analysis. B. Table of principal components and P Values C. Scatter graphs plotting the first 3 Principal Components (PC1,PC2,PC3) accounting for >95% of variance, with tumour data plotted in red, non-tumour data in green

PCA fed LDA analysis was performed using all 12 principal components which gave a training performance of 90.5%. LDA LOOCV gave a sensitivity of 90.2%, specificity 90.5%, with overall accuracy of 90.3% (Figure 4-28).

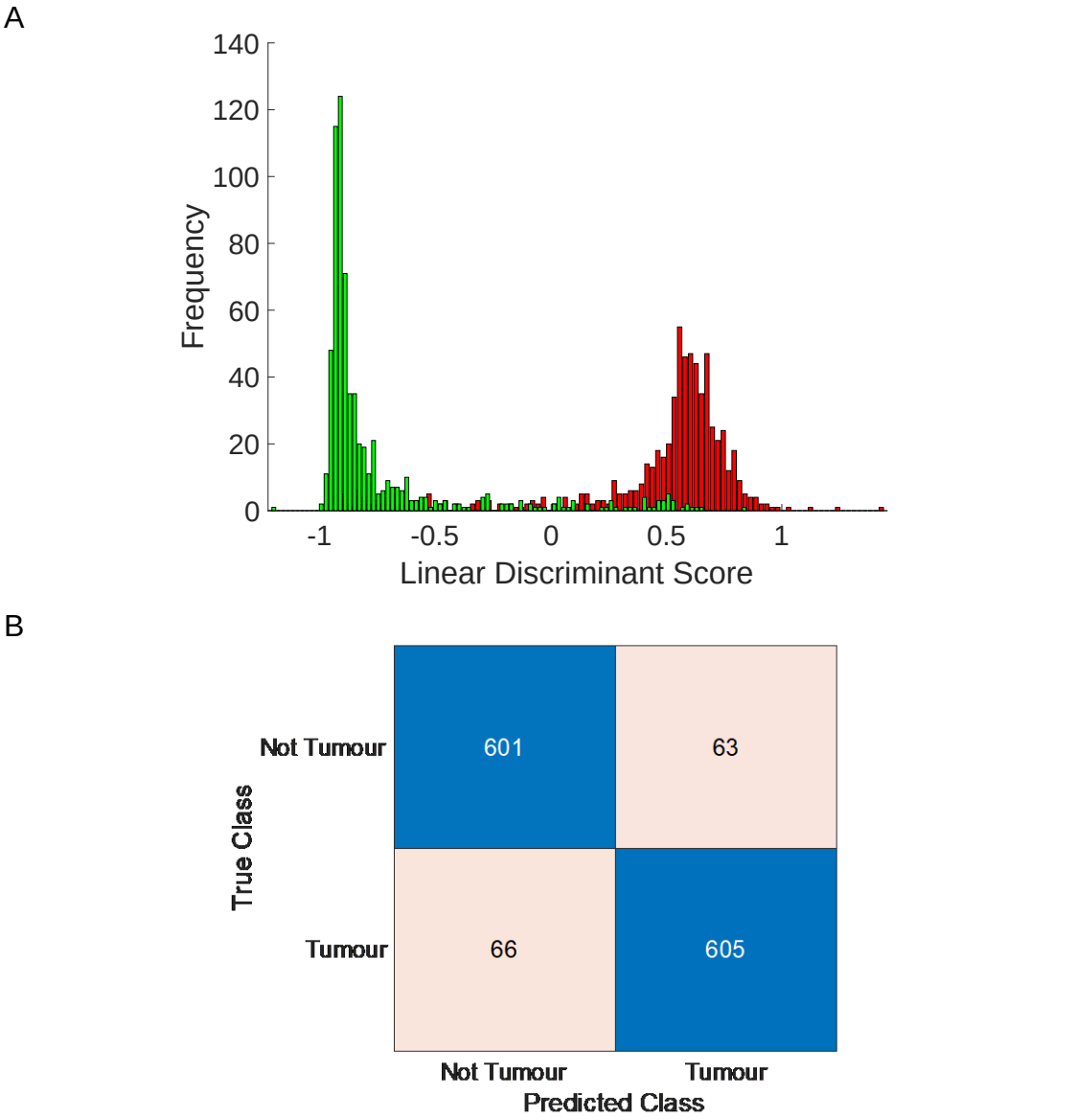
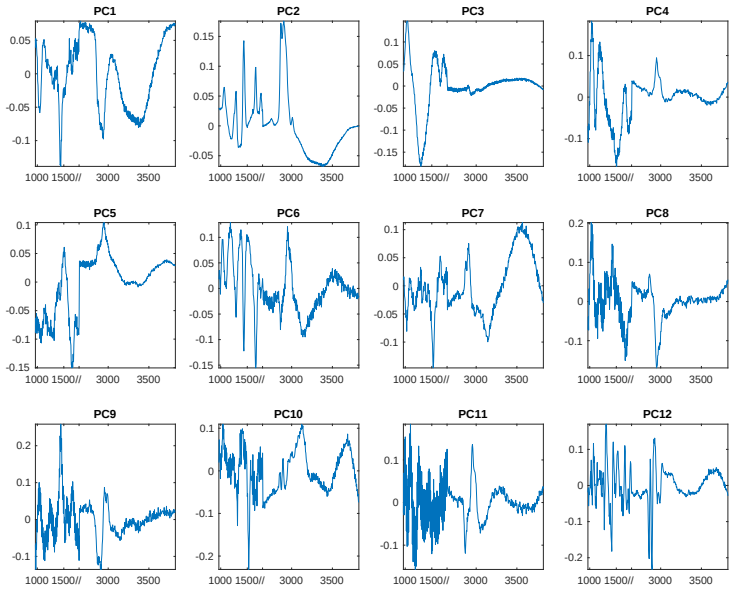


Figure 4-28 PCA fed LDA analysis of all concatenated spectral data (n=1335)

A PCA fed LDA analysis histogram of scores versus frequency, with tumour data plotted in red, non- tumour data in green. B. confusion matrix prediction using LDA LOOCV

The mean concatenated spectrum for each specimen was calculated and underwent PCA analysis. The 2 significant PC's corresponded to protein (PC1) and lipid (PC2) signals, which had significant scores ($P<0.01$). These two principal components accounted for 63.86% (PC1) and 32.7% (PC2) of variance, with the remaining 10PC's accounting for the remaining <4% of variance. Plotting the scores from the PC1 and PC2 that accounted for > 95% variance demonstrated they achieved good separation (Figure 4-29).

A



B

PC 1	PC 2	PC 3	PC 4
$1.67X10^{-62}$	$4.70X10^{-65}$	0.556	0.397
PC 5	**PC 6**	**PC 7**	**PC 8**
0.858	0.908	0.774	0.915
PC 9	**PC 10**	**PC 11**	**PC 12**
0.763	0.267	0.741	0.574

C

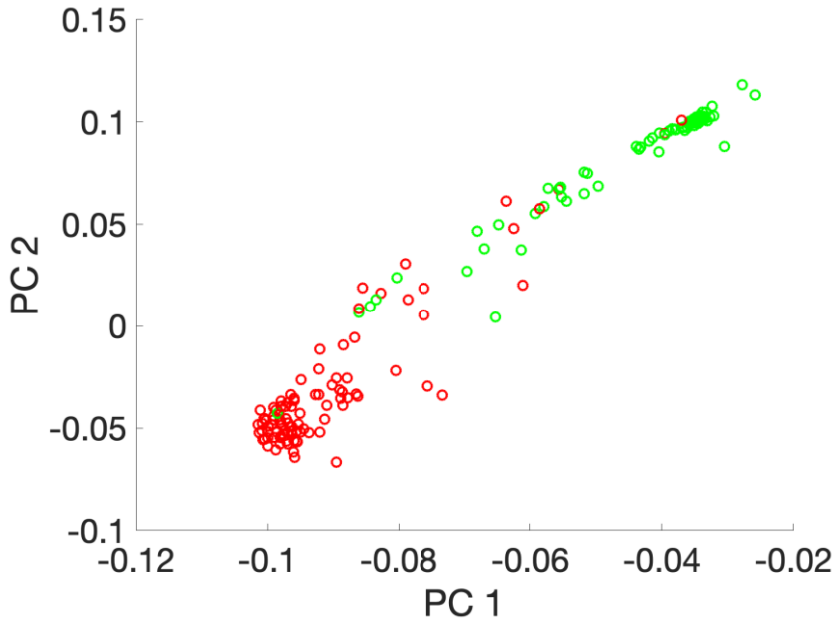


Figure 4-29 Principal Component Analysis of concatenated spectra of mean spectra (n=192)

A. Principal Components (12) identified in analysis B. Table of principal component scores and corresponding P Values C. Graph plotting the first 2 significant Principal Components (PC1 and PC2) that accounted for >95% variance, with tumour data plotted in red, non- tumour data in green

PCA fed LDA analysis was performed, which gave a training performance of 94.3%, with LOOCV the prediction gave an overall accuracy of 92.7%, sensitivity 93.8% and specificity of 91.7% (Figure 4-30).

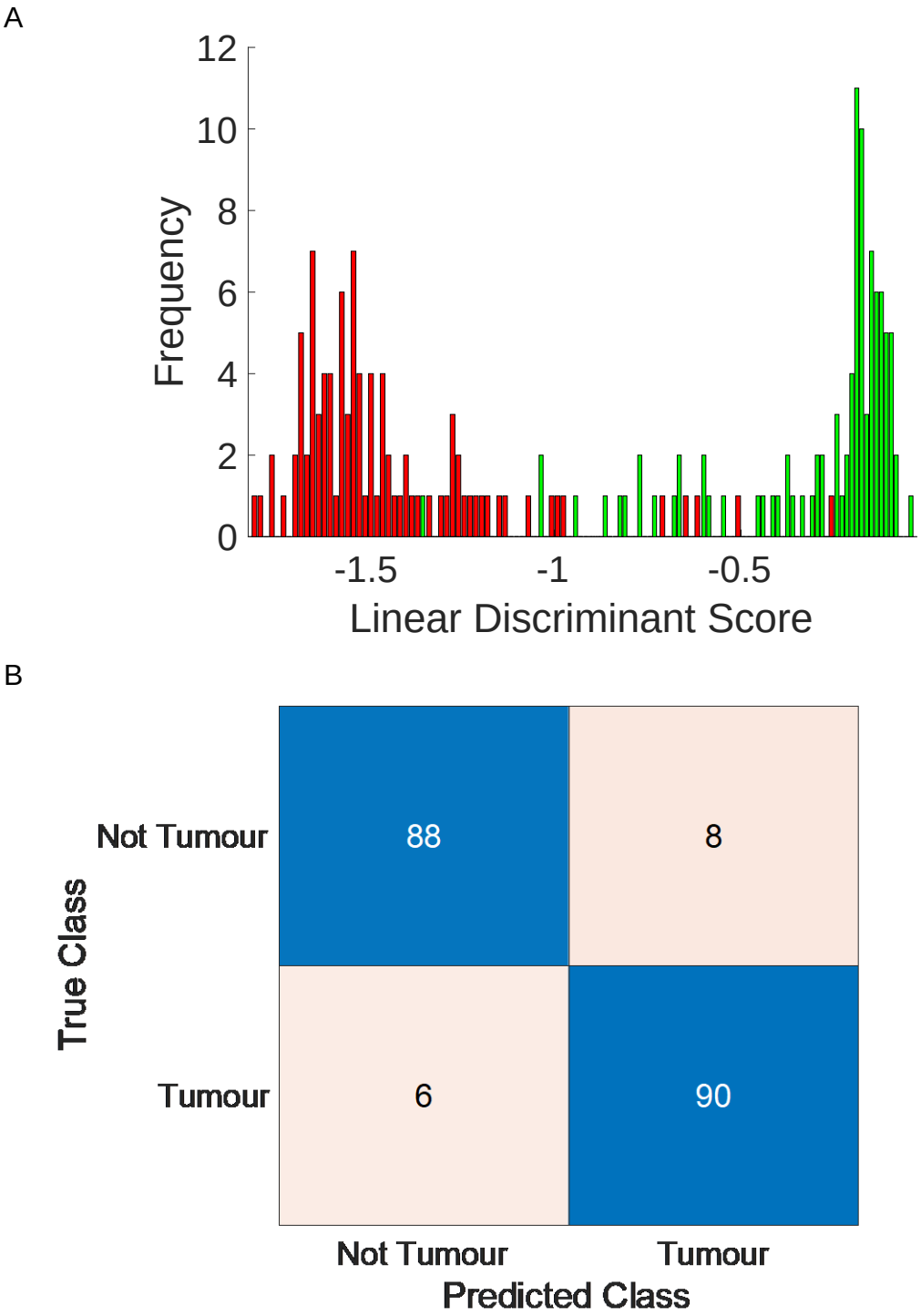


Figure 4-30 Analysis results of mean concatenated spectra for each specimen (n=192)

A. Histogram of PCA fed LDA scores versus frequency, with tumour data plotted in red, non- tumour data in green. B. Confusion matrix of LDA LOOCV prediction results

4.3.2.4 Summary statistics for NP3 ability to differentiate between tumour and non-tumour breast tissue

Spectral Region		**Analysis method**	**Sens.**	**Spec.**	**Overall Accuracy**
HWN	All spectra	PCA/LDA	90.2	90.5	90.3
	Mean spectra		93.8	92.7	93.2
FP	All spectra		90.2	90.5	90.3
	Mean spectra		91.7	93.8	92.7
Concatenate	All spectra		90.2	90.5	90.3
	Mean spectra		93.8	91.7	92.7
W/TR	All spectra	Binomial Logistic regression†	92.7	86.6	89.7
		Classification trees	87.5	91.7	89.6
	Mean spectra	Binomial Logistic regression†	94.8	89.6	92.2
		Classification trees	92.7	90.6	91.7
	Representative spectrum/sample	Voting threshold 1	87.5	95.8	91.7
W/TAR Difference	All spectra	cut off 0.1	93.6	85.4	89.5
	All spectra	cut off 0.33†	87.0	93.2	90.1

Table 4-7 Summary statistics table of best performing spectral analysis and water/total area ratio analysis HWN – High Wavenumber region; FP – Fingerprint; Concatenate – HWN +FP spectra stitched together; W/TR – Water/ total area ratio; PCA/LDA – PCA fed LDA analysis; Sens. – sensitivity, Spec. - Specificity † - binomial logistic regression with k-folds 5 cross validation

4.3.2.5 Spectral analysis to differentiate between pathology subtypes

The group of tumour specimens contained a wide range of different pathology subtypes. Spectral analysis was performed to investigate if Raman measurements had the ability to provide specific pathological diagnosis of the tumour.

Using E(strogen) R(eceptor) status as a representative example, the spectra obtained from all tumour samples showed little difference between those tumour samples that were ER+ versus those that were ER- by the sequence of analysis of looking at the water/total area ratio (Figure 4-31 A), and spectral analysis of the concatenated data with PCA (Figure 4-31 B). The findings were similar in spectral analysis of all pathological subtypes.

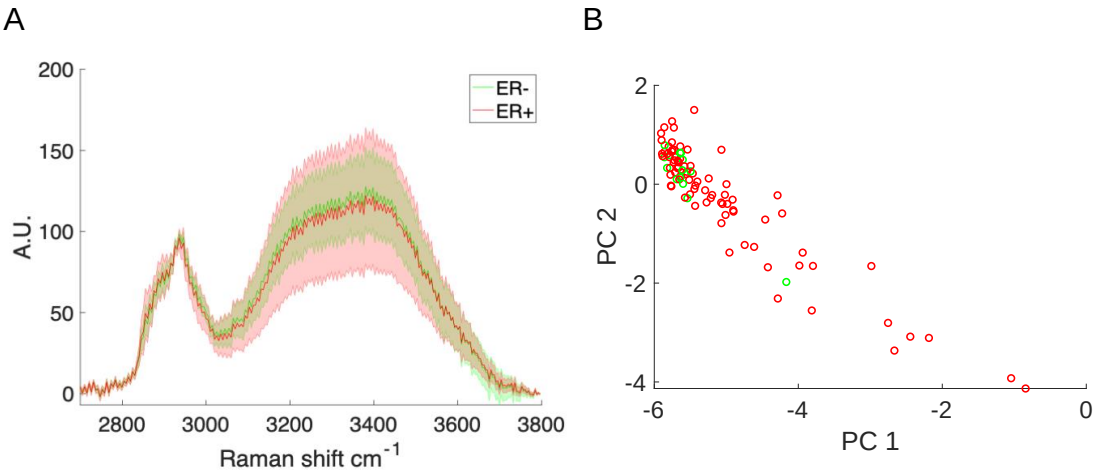


Figure 4-31 Results of mean tumour specimen (n=96) spectral analysis of ER status (used as a representative example)

A. Mean HWN spectral data of all tumour specimens according to ER status. Data line is mean of spectra from each sample of that pathology subtype; Red – ER+ (n=72), Green – ER- (n=24). Shaded area either side in the same colour is +/- 1 S.D. Spectra were baselined using 3rd order polynomial and normalised to 2985cm^{-1} peak. B. Scatter graph of first 2 principal components (PC1 and PC2) scores derived from PCA analysis of concatenated spectral data. Each circle represents a specimen mean data, Red – ER+, Green – ER-.

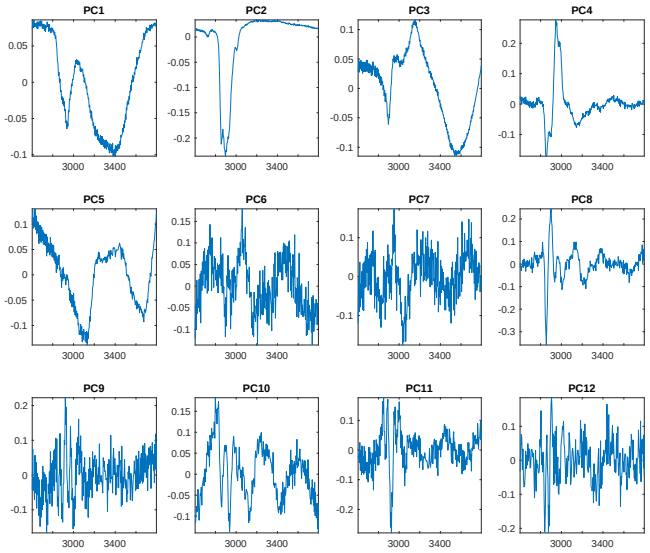
In the analysis of tumour versus non-tumour differentiation, it was the mean data of the specimens that had the highest diagnostic accuracy of the water/total area ratio, and the mean data of the concatenated spectral data. Therefore, in the analysis of the different pathology types in tumour specimens, these data only were analysed, as if there was poor diagnostic accuracy in differentiating pathology types using these data, the analysis of all spectra would be presumed

to be worse. The tumour specimens were analysed to investigate if there were any spectral differences according to pathology groups using the water/total area ratio and a PCA fed LDA analysis.

Spectral analysis with PCA was performed according to each pathological subtype. It was performed using the HWN spectral region (Figure 4-32) and the fingerprint spectral region (Figure 4-33).

The PCA scores did not reach statistical significance in either the HWN or FP data for any of the pathological subtypes apart from carcinoma type. The principal components obviously relating to the protein, lipid and water peak in both the HWN and FP region (PC1 and PC 2) did not reach statistical significance ($P>0.01$) in the majority of pathological subtypes, suggesting that these peaks were not different between pathological subtypes (unlike in tumour versus non-tumour differentiation). Although, in the HWN region PC2 (that accounted 4% of variance) and PC4 (accounted for 0.2% variance) were significant ($P<0.01$) for carcinoma type. These both relate to lipid signal, suggesting there may be a difference in lipid signal between some carcinoma type. Also, in the FP region PC 5 (accounted for 0.5% variance) and PC6 (accounted for 0.2% variance) were significant ($P<0.01$) for carcinoma type. These both relate to the amide III region of protein, suggesting that difference in protein signal can differentiate between carcinoma type. However, conclusions of biochemical differences between carcinoma types are limited based on these results, as the significant principal components account for very little of the variance between specimens.

A



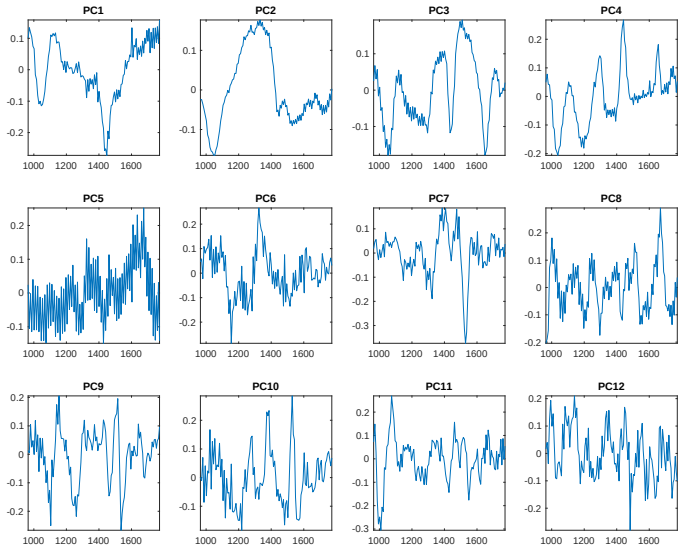
B

Pathological Subtype	**Principal Component**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ER	0.09	0.30	0.02	0.07	0.43	0.40	0.68	0.30	0.17	0.64	0.23	0.62
HER2	0.62	0.41	0.71	0.54	0.84	0.17	0.49	0.42	0.43	0.18	0.16	0.76
Carcinoma Type	0.03	<0.01	0.57	<0.01	0.89	0.98	0.36	0.99	0.31	0.17	0.14	0.26
Carcinoma Grade	0.95	0.92	0.72	0.14	0.84	0.22	0.62	0.75	0.03	0.61	0.76	0.87
Assoc. w/DCIS	0.08	0.26	0.07	0.57	0.54	0.21	0.15	0.04	0.81	0.38	0.25	0.42
Lymph Node status	0.56	0.43	0.92	0.87	0.81	0.75	0.25	0.24	0.10	0.56	0.65	0.43

Figure 4-32 Principal component analysis of mean HWN tumour spectra according to pathological subtype classification (n=96)

A Principal components of mean HWN spectra of all tumour specimens. B Summary table of Principal component score P values according to pathological subtype classification. P values of principal component scores from one-way ANOVA analysis.

A



B

Pathological Subtype	Principal Component											
	1	2	3	4	5	6	7	8	9	10	11	12
ER	0.15	0.16	0.69	0.79	0.44	0.58	0.64	0.31	0.03	0.07	0.52	0.28
HER2	0.36	0.70	0.40	0.19	0.06	0.37	0.88	0.99	0.32	0.89	0.80	0.87
Carcinoma Type	0.60	0.05	0.02	0.73	<0.01	<0.01	0.18	0.93	0.96	0.81	0.49	0.92
Carcinoma Grade	0.69	0.80	0.48	0.11	0.09	0.71	0.80	0.91	0.10	0.24	0.09	0.47
Assoc w/DCIS	0.03	0.05	0.56	0.46	0.73	0.37	0.47	0.65	0.55	0.49	0.43	0.67
Lymph Node status	0.85	0.36	0.36	0.91	0.13	0.10	0.19	0.83	0.10	0.87	0.55	0.64

Figure 4-33 Principal component analysis of mean FP tumour spectra specimens according to pathological subtype classification (n=96)

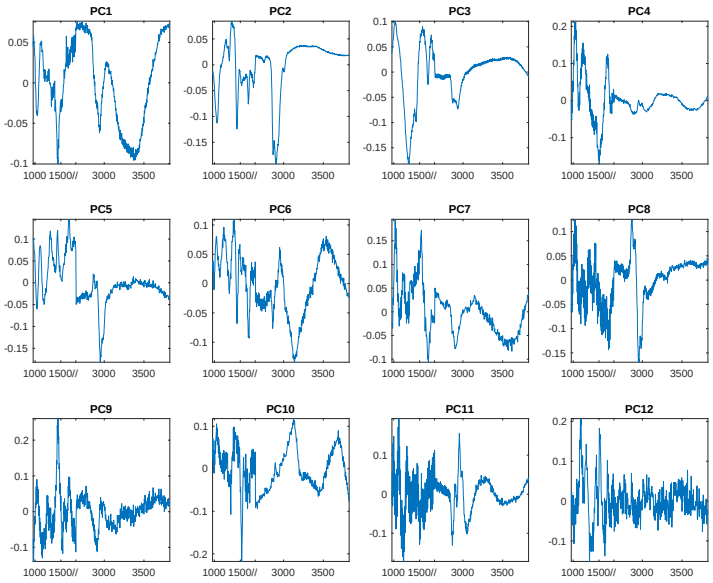
A Principal components of mean FP spectra of all tumour specimens. B Summary table of Principal component score P values according to pathological subtype classification. P values of principal component scores from one-way ANOVA analysis.

Given the results from spectral analysis of the HWN and FP region suggest that these individual regions cannot differentiate between tumour pathological subtypes, the concatenated spectral data was analysed to investigate if this improved accuracy (Figure 4-34). The only significant result was in the different carcinoma types, where the 'carcinoma type' classification did reveal a single significantly different score in PC 2 and PC5 ($P<0.01$). Both PC2 (accounting for 4.5% of variance) and PC5 (accounting for 0.36% of variance) relate to the lipid peak in both the FP and HWN region (by separate spectral peak analysis not displayed). This suggests that spectral features of lipid differentiate between different carcinoma types. Again, these significant components only account for a small degree of variance between spectra.

The PCA scores were used in a PCA fed LDA analysis for all pathological subtypes using the three different spectral regions (HWN, FP and Concatenated), to investigate if spectral differences could predict pathological subgroups- the results are in Table 4-8. The results show that for the HWN region and FP region results were similar between spectral regions – out of 6 subgroups tested, three had better prediction results from the HWN region spectra than the FP region, and vice versa. The concatenated data compared to the individual regions was very similar in the prediction ability. All prediction accuracies were between 40% (for lymph node status prediction using FP region) up to 78% (for ER prediction using FP region.

Overall the results are poor and suggest that with the NP3 system, neither HWN, FP or concatenated spectral analysis can accurately differentiate between tumour specimen pathology subtypes.

A



B

Pathological Subtype	Principal Component											
	1	2	3	4	5	6	7	8	9	10	11	12
ER	0.08	0.17	0.47	0.77	0.88	0.01	0.62	0.54	0.12	0.60	0.29	0.76
HER2	0.55	0.38	0.88	0.72	0.18	0.76	0.16	0.25	0.65	0.36	0.33	0.45
Carcinoma Type	0.04	<0.01	0.60	0.24	<0.01	0.08	0.73	0.70	0.60	0.36	0.91	0.54
Carcinoma Grade	0.93	0.92	0.86	0.15	0.47	0.51	0.07	0.07	0.31	0.98	0.88	0.65
Assoc w/DCIS	0.06	0.09	0.23	0.74	0.23	0.37	0.77	0.85	0.17	0.76	0.42	0.48
Lymph Node status	0.62	0.30	0.81	0.48	0.88	0.45	0.32	0.62	0.07	0.64	0.08	0.59

Figure 4-34 Principal component analysis of mean concatenated tumour spectra according to pathological subtype classification (n=96)

A Principal components of mean concatenated spectra of all tumour specimens. B Summary table of Principal component score P values according to pathological subtype classification. P values of principal component scores from one-way ANOVA analysis.

Pathological Subtype	PCA fed LDA accuracy (%)		
	HWN region	FP region	Concatenated
ER	76.0	78.1	77.1
HER2	63.5	61.4	66.7
Carcinoma Type	58.3	61.5	61.5
Carcinoma Grade	59.4	64.5	61.5
Assoc w/DCIS	73.9	65.6	66.7
Lymph Node status	51.0	40.6	41.7

Table 4-8 Results of ability of the three different spectral regions to predict pathological subtype by PCA fed LDA analysis performed on mean spectral data (n=96)

The analysis of the concatenated spectral data was compared to the analysis of the ability of water/total area ratio to differentiate between different pathological subtypes (Table 4-9). The mean water/total area ratio was calculated for each pathological subtype and compared between the other subtypes with a student t-test (for dichotomous groups) or one-way ANOVA (for more than 2 groups). The outcomes are displayed in Table 4-9, which shows that there were no water/total area ratios that were significantly different to another. The mean water/total area ratio was close to reaching significance between carcinoma types (P=0.0185, one-way ANOVA), and on comparison of multiple means the only difference was the ratio between ductal carcinoma (mean ratio 0.78; S.D. 0.11 (n=64)) versus lobular carcinoma (0.70; S.D.0.16 (n=21)) which did not reach statistical significance (P=0.03). This suggests that areas of ductal carcinoma tissue may have a higher water/total area ratio than lobular carcinoma tissue. Using a multiclass error-correcting output codes (ECOC) model based on the water/total area ratio to predict carcinoma type had an AUC of 0.64, which suggests that using the water/total area ratio to predict different carcinoma types performed poorly.

The consistent significant finding in both the spectral analysis of all three regions, and the water/total area ratio analysis, was the difference in carcinoma type. The water/total area ratio suggested there was a difference between ductal and lobular carcinoma. Spectral analysis also suggested differences in protein and lipid peaks between carcinoma types which is concordant with the known

histopathological differences between ductal and lobular tumours (lobular having more fatty tissue, which is explored later in this chapter in our histological study). To investigate this a PCA fed LDA sub-analysis was performed, grouping all carcinoma types that were not lobular together, and compared all carcinoma types versus lobular carcinoma (not displayed). The PCA showed a significant difference in scores of PC1 and PC2, that accounted for changes in protein, lipid and water changes. The PCA fed LDA analysis had a prediction accuracy of 81.2%, which is higher than when the carcinoma types were not grouped (61%). This suggests that there are biochemical differences between ductal and lobular carcinoma (these are investigated and discussed in the histological study), which can be detected with our Raman system. However, both models (water/total area ratio and PCA/LDA) performed poorly when using the spectral data to predict between carcinoma types suggesting the spectral data could not accurately differentiate between carcinoma type. It may be that this study was under powered to detect differences in carcinoma type (the number of lobular carcinoma was 21). It may be that with a larger data set, these differences may reach greater significance, and allow prediction models to differentiate between different carcinoma types.

Analysis of the other tumour characteristics of ER and HER2 status, carcinoma grade, whether the tumour was associated with DCIS and axillary lymph node status revealed no significant difference in comparison of the mean ratios ($P>0.01$), and demonstrated the AUC of the binomial logistic regression model with cross validation was generally poor ranging between 0.47 – 0.54, which is similar to random chance of classification. Analysing the concatenated spectra data with PCA fed LDA analysis according to the pathological subtypes revealed a poor ability to accurately predict pathological subtype with the accuracy ranging from 45 – 77%.

These results suggest that the spectra obtained in these experiments cannot accurately differentiate between different pathological subtypes of breast tumour.

Pathology subtype	Status	Mean ratio[+]	Comparison of mean P value[+°]	AUC[+]	Spectral accuracy[†]
ER	+	0.76	0.19	0.54	77%
	-	0.81			
HER2	+	0.75	0.52	0.49	66%
	-	0.77			
Carcinoma Type	Ductal	0.78	0.0185	0.64	61%
	Lobular	0.70			
	Mixed	0.73			
	Mucinous	0.87			
Carcinoma Grade	1	0.83	0.87	0.48	61%
	2	0.77			
	3	0.77			
Associated w/ DCIS	+	0.76	0.13	0.50	66%
	-	0.80			
Axillary lymph node status	+	0.75	0.15	0.47	41%
	-	0.79			

Table 4-9 Table demonstrating analysis of mean spectra of tumour specimens (n=96) of pathology subtypes

AUC was calculated from binomial logistic regression in dichotomous classifiers, and a multiclass ECOC model for multiple classifiers using mean ratio values and k-folds 10 cross validation

Spectral accuracy data was from the mean spectral data, with a 12 component PCA fed LDA model, and the resultant accuracy of the LDA model is quoted

[+] - denotes values derived from spectral data that had been baselined with 3rd order polynomial and the mean water/total area ratio calculated

[†] - denotes values derived from mean concatenated spectral data

°- for comparison of two means, unpaired t-test, for comparison of multiple means, one-way ANOVA

4.3.2.6 Ability of Raman to differentiate DCIS only from non-tumour tissue

The majority of the tumour samples were also associated with DCIS (77%), therefore, when measuring a 'tumour' sample, the measurement could be of the invasive element of the tumour, or the associated DCIS. Therefore, in the analysis of tumour versus non – tumour above, the samples that contained DCIS only, and no invasive tumour (n=3) were included as 'tumour' samples. DCIS is responsible for the majority of positive margins [187], therefore the ability to differentiate DCIS only from non-tumour tissue is important in furthering the goal of being able to provide IMA.

To understand if the NP3 system, and the data processing techniques, can differentiate DCIS from non-tumour tissue, the samples from the patients with DCIS only tissue (n=3) underwent a separate analysis.

The HWN spectra of DCIS only tissue compared to the matched non-tumour tissue demonstrated obvious differences, with DCIS only tissue having strong protein and water peaks, and non-tumour tissue showing lipid peaks with little water signal (Figure 4-35). The water/total area ratio was calculated which showed a difference between DCIS only tissue (mean water/total area ratio- 0.83; SD- 0.019) and non-tumour tissue (mean water/total area ratio – 0.22; SD 0.19) which did not reach statistical significance (P= 0.035), it is possible the small number of samples meant the 99% confidence level could not be attained.

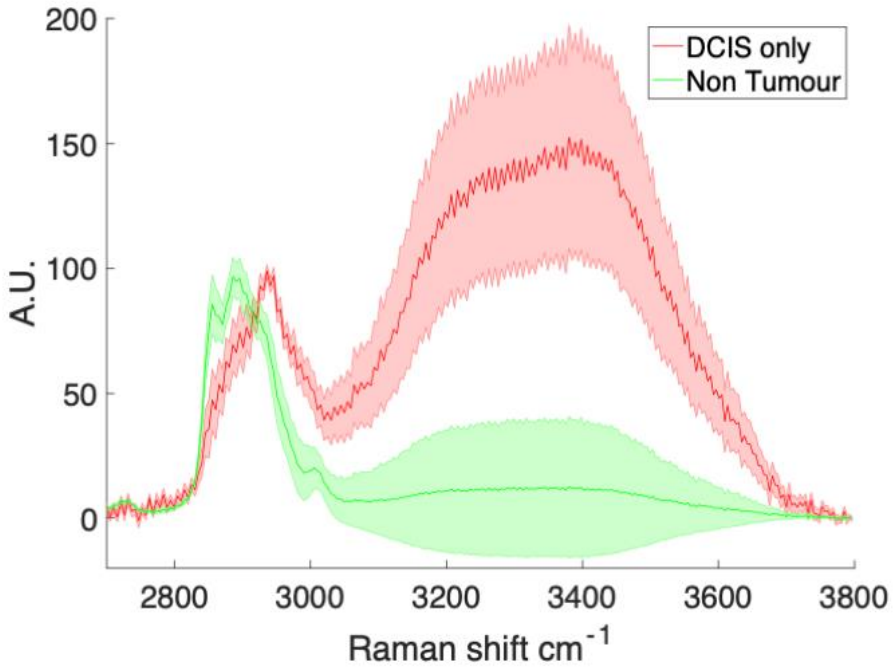


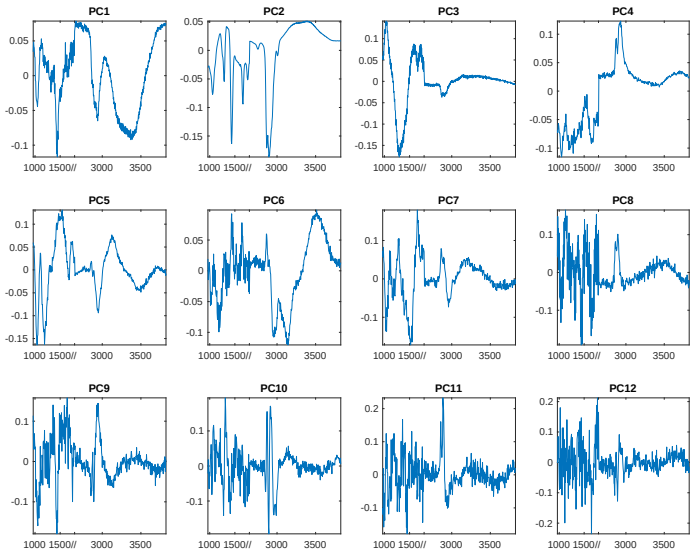
Figure 4-35 Mean HWN spectra of all specimens (spectra n=42, patients n=3) according to diagnosis demonstrating changes in water content between DCIS only and non-tumour specimen spectra
Data is mean of spectra from all DCIS only specimen spectra (n=21) and the matched non- tumour specimen spectra (n=21), fluorescent baseline subtraction with a 3rd order polynomial, and normalised to the CH peak at 2935cm^{-1}. Shading either side of solid line in same colour is +/- 1 S.D. Red spectra=measurement from tumour specimen, Green spectra = measurement from non-tumour specimen

The water/total area ratio was used in a binomial logistic regression to construct a ROC curve and obtain the optimal threshold with k-folds 5 cross validation for classifying the specimen as DCIS or non- tumour tissue. All spectral measurements (n=42) were used. The model showed that with a water/total area ratio threshold of 0.76, it achieved an AUC of 0.96, sensitivity of 100% and specificity of 90.5%. With such a small number of readings, and all spectra obtained from just 3 patients, this should be interpreted with caution, but it does suggest that the water/total area ratio can differentiate between DCIS only and non-tumour tissue and may be used for diagnostic prediction.

Spectral analysis was performed on the concatenated spectra of all measurements (n=42) (Figure 4-36). 97% of the variance was accounted for by PC1 (61.8% of variance), PC 2 (31.16% of variance) and PC 3 (4.7% of variance) of which PC 1 and PC 2 scores achieved statistical significance (P<0.01), these PC's related to protein and water (PC1) and lipid (PC2 and PC3) spectra and plotting these two PC's gave excellent separation, suggesting that the

differentiation between DCIS only and non-tumour tissue is based on the same spectral differences as between tumour and non-tumour tissue.

A



B

PC 1	PC 2	PC 3	PC 4
3.3×10^{-19}	3.3×10^{-17}	0.19	0.294
PC 5	**PC 6**	**PC 7**	**PC 8**
0.56	0.54	0.97	0.94
PC 9	**PC 10**	**PC 11**	**PC 12**
0.88	0.43	0.60	0.95

C

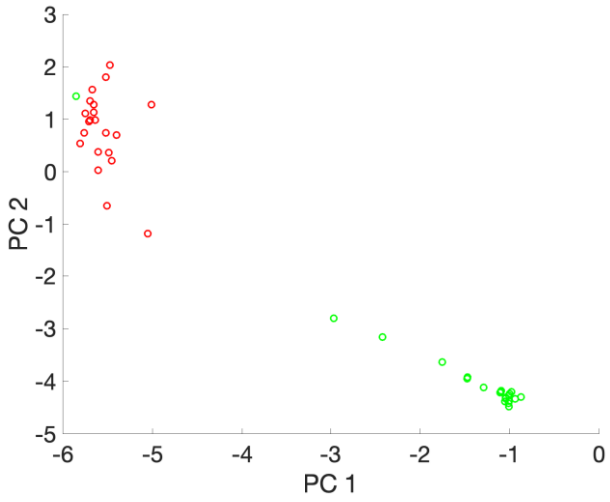


Figure 4-36 Principal Component Analysis of concatenated spectral data for patients with DCIS only (n=42)

A. Principal Components (12 in total) identified in analysis B. Table of principal components and corresponding P Values C. Graph plotting the 2 significant Principal Components (PC1 and PC2), with tumour scores plotted in red, non-tumour scores in green

PCA fed LDA analysis with LOOCV gave a 95.2% accuracy, with 95.2% sensitivity and 95.2% specificity. As the numbers of spectra are so small it does not give a statistically robust outcome, but it does suggest that the spectral features of DCIS only tissue from Raman measurements and our data analysis protocol can differentiate between DCIS only and non-tumour breast tissue.

4.3.2.7 Spectral analysis of non- tumour specimens to differentiate between physiological variation

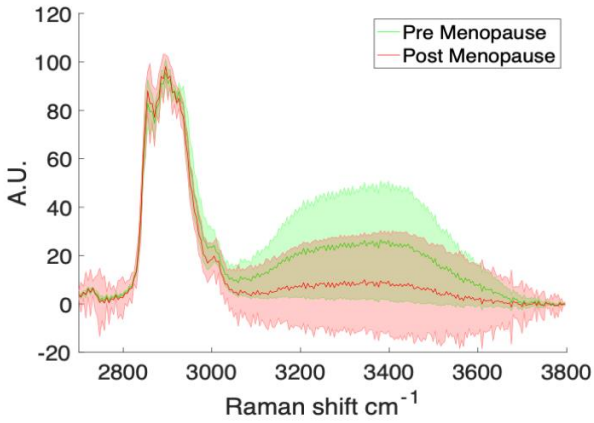
Spectra from non-tumour specimens were analysed according to patient characteristic information that were collected from patients which may be hypothesised to lead to variation in normal breast tissue composition between patients based on the literature – these were age and menopausal status [280, 306]. Patients were classed according to two age brackets – with an age cut off of <60 and <50, and menopausal status (pre and post), the summary results table is seen in Table 4-10.

Classifying spectra obtained from non-tumour specimens according to menopausal status found that the water/total area ratio was higher in the pre-menopausal (water/total area ratio 0.3374; SD 0.21)(n=73) patients compared to the post – menopausal (0.1858; SD 0.19)(n=16) patients, which reached statistical significance (P=0.006; two tailed t-test). In Figure 4-37 B, the principal components of the PCA analysis of all the non-tumour specimen spectra can be seen. PC 1 which accounted for 89.8% of variance did not reach statistical significance (P>0.01), however PC 2 (accounting for 8.2% of variance) reached statistical significance (P<0.01) and relates to the protein and water peak in the HWN region, suggesting that there are differences in these areas between the groups of patients.

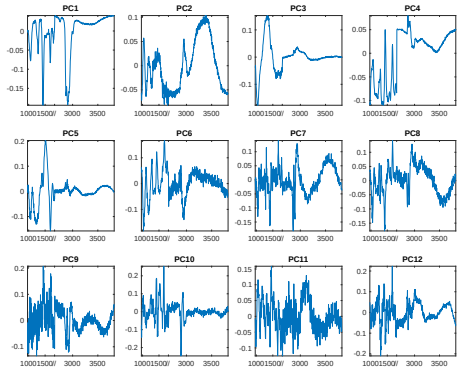
Using the same classification of menopausal status to classify the specimen concatenated spectral data, PCA fed LDA achieved an accuracy of 84%, suggesting there were differences in the spectral data between patients that were pre- compared to post- menopausal that could allow prediction. Due to the significance of the LDA analysis, a LOOCV analysis was performed to assess the accuracy of prediction on spectral data using LDA, and a sensitivity of 80.2%, specificity of 56.2% and overall accuracy of 76.4% was achieved. This suggests

that using the PCA fed LDA analysis data that did not form the training set is not particularly accurate at predicting whether non-tumour specimens were from pre or post-menopausal patients. Regardless of this, the results suggest there are differences in the non-tumour breast tissue between pre and post-menopausal patients, and due to the difference in the water/total area ratio, and the P values of the principal components relating to the HWN region, it suggests that normal breast tissue of pre-menopausal women have a higher water content.

A



B



C

PC 1	PC 2	PC 3	PC 4
0.04	0.008	0.60	0.82
PC 5	**PC 6**	**PC 7**	**PC 8**
0.88	0.41	0.74	0.13
PC 9	**PC 10**	**PC 11**	**PC 12**
0.40	0.09	0.36	0.01

D

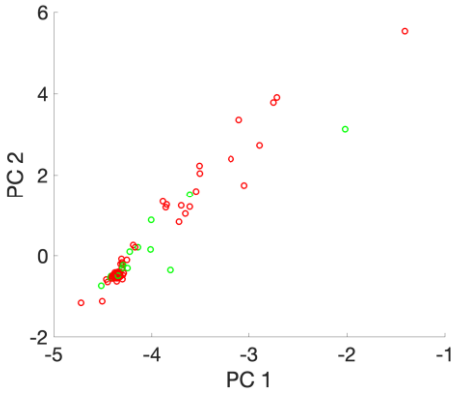


Figure 4-37 Results of mean spectral analysis of non-tumour breast specimens classified according to menopausal status (n=89)

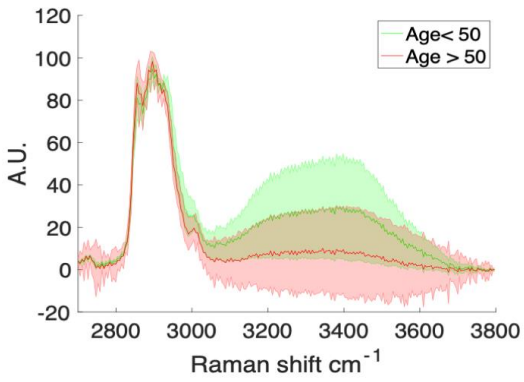
A. Mean data of all tumour specimens according to menopausal status. Data line is mean of spectra from each sample of that pathology subtype; Red – post- menopausal (n=73), Green – pre-menopausal (n=16). Shaded area either side in the same colour is +/- 1 S.D. Spectra were baselined using 3rd order polynomial and normalised to 2935cm^{-1} peak. B. Principal components of mean concatenated spectra of all tumour specimen (n=89). PCA analysis performed on mean concatenated spectra. C. P values of principal component scores according to menopausal status. D. Scatter graph of PC1 and PC2 scores that accounted for >97% of variance, PC2 also had statistical significance. Each circle represents a specimen mean data, Red – post-menopausal, Green – pre-menopausal.

Spectral data was then classified according to age. Using the age cut off of 60 to classify the water/total area ratio, the non-tumour specimens of those <60 had a higher water/total area ratio (mean water/total area ratio 0.31; SD 0.21)(n=27) compared to those patients >60 (mean water/total area ratio 0.18; SD 0.19) (n=68), this reached statistical significance (P= 0.0063; two tailed t-test). Analysis of concatenated spectral data using a PCA fed LDA analysis showed a model accuracy of 78.9%, and with cross validation with LOOCV it achieved an accuracy of 75.8%, with a sensitivity of 80.9%, and specificity of 63% showing poor ability for prediction.

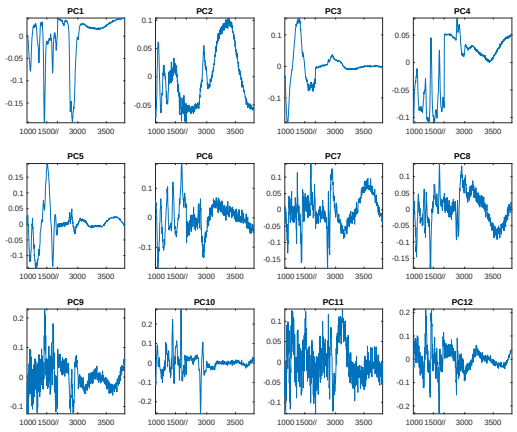
Using the age cut off of 50 to classify the water/total area ratio results showed a greater difference between the groups. Figure 4-38 shows the results of the analysis, and in figure A the difference in the water peak between the two groups can be seen, with the younger group having a higher water peak. The non-tumour specimens of those <50 had a higher water/total area ratio (mean water/total area ratio 0.37; SD 0.21) (n=16) compared to those patients > 50 (mean water/total area ratio 0.19; SD 0.19) (n=79), this reached statistical significance (P=0.00063; two tailed t-test). Despite this highly significant P value, using binomial logistic regression of water/total area ratio and age of 50 as a classifier, the model achieved poor diagnostic ability with an accuracy of 69%. Analysis of the concatenated spectral data using a PCA fed LDA analysis showed a model accuracy of 85.3%. The principal components PC 1 (accounting for 89.5% of variance) and PC2 accounting for 8.51% of variance) gained statistical significance in their scores with age as the classifier (P<0.01), related to the lipid (PC 1) and protein and water peak (PC 2) (Figure 4-38). The PCA scores underwent LOOCV LDA analysis to validate the findings, and found a sensitivity of 82.3%, specificity of 56.2% and accuracy 77.9%.

These results demonstrate that there is a difference in the normal breast tissue composition according to the age of the patient. Younger patients have a higher water/total area ratio, suggesting the breast tissue has a higher water content. The difference is more obvious in using the younger age cut off of 50, with a highly significant different water/total area ratio, however the accuracy of being able to predict the classifier is modest.

A



B



C

PC 1	PC 2	PC 3	PC 4
0.003	0.0013	0.082	0.1236
PC 5	**PC 6**	**PC 7**	**PC 8**
0.9453	0.2955	0.6142	0.0173
PC 9	**PC 10**	**PC 11**	**PC 12**
0.4727	0.3690	0.1128	0.1465

D

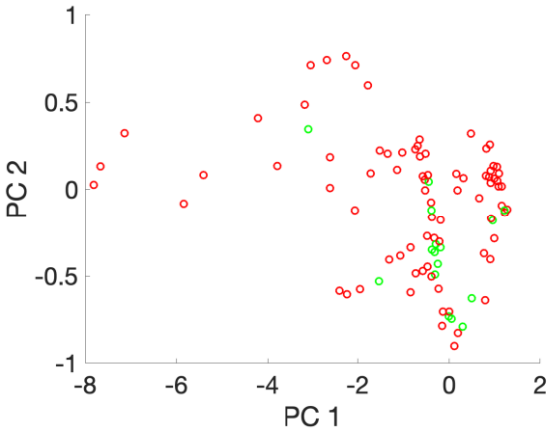


Figure 4-38 Results of spectral analysis of non-tumour breast specimens classified according to Age (cut off 50) (n=95)

A. Mean data of all tumour specimens according to Age (cut off 50). Data line is mean of spectra from each sample of that pathology subtype; Red – Age >50 (n=79), Green – Age <50 (n=16). Shaded area either side in the same colour is +/- 1 S.D. Spectra were baselined using 3rd order polynomial and normalised to 2985cm^{-1} peak. B. Principal components of mean Concatenated spectra of all tumour specimen (n=95). C. P values of principal component scores according to age (cut off 50). D. scatter graph of the significant principal components (PC1 and, PC2) scores that also accounted for >95% of variance. Each circle represents a specimen mean data, Red - Age >50, Green - Age <50.

The importance of these findings summarised in Table 4-10 is not in the prediction ability of the models, but in the difference in water/total area ratios. This demonstrates that there are significant physiological variations between the normal, non-tumour tissue of patients according to menopausal status and age. In the use of the HWN spectrum, or the water/total area ratio, to diagnose breast cancer it must be considered that the baseline 'normal' to which the tumour spectra are being compared is a fluctuating baseline that changes according to the physiological characteristics of the patient.

Patient characteristic	Status	Mean ratio^{+}	Comparison of mean P value$^{+\circ}$	AUC^{+}	Spectral accuracy†
Menopausal status	Pre	0.3374	0.006	0.65	84%
	Post	0.1858			
Age - 60	<60	0.31	0.0063	0.65	78.9%
	>60	0.18			
Age – 50	<50	0.37	0.00063	0.69	85.3%
	>50	0.19			

Table 4-10 Summary results table of analysis of non-tumour specimens mean spectra classified by physiological characteristics

AUC was calculated from binomial logistic regression using mean ratio values and k-folds 10 cross validation

Spectral accuracy data was from the mean spectral data, with a 12 component PCA fed LDA model, and the resultant accuracy of the LDA model is quoted

$^{+}$ - denotes values derived from spectral data that had been baselined with 3rd order polynomial and the mean water/total area ratio calculated

† - denotes values derived from mean concatenated spectral data

°- comparison of two means, unpaired t-test

4.3.3 Estimation of water content by dehydration

Tumour Specimens

Five tumour specimens underwent serial dehydration, of these five, two samples and the corresponding Raman data were discarded due to the samples becoming so small through dehydration that accurate Raman measurement or weighing was not possible. Therefore, the weights and corresponding Raman data of three specimens were included for analysis.

The starting weight of the three specimens were 0.167 g, 0.083 g and 0.089 g, the dry weight was 0.041, 0.019 and 0.022 g, giving a total weight loss of 75%, 77% and 75% of the starting weight respectively. Corresponding HWN Raman spectra over the period of dehydration showed visual changes in the spectra with a decrease in the water peak (in spectra normalised to the protein peak), as shown in Figure 4-39 A. This suggests that the weight loss was secondary to dehydration, and that the HWN Raman measurements could measure these changes.

The water/total area ratio was calculated for all acquired Raman spectra. These were plotted against the known water fraction of the specimen at the time the Raman spectra was acquired, the results are shown in Figure 4-39 B-E. In graph E the dehydration curve of the Raman spectra is comparable between all 3 spectra, suggesting that the changes in water/total area ratio with water fraction of tumour breast tissue is repeatable, and equivalent between different patients.

These results demonstrate that changes in water concentration in tumour containing breast tissue can be measured, and to a certain extent, quantified by HWN Raman spectroscopy, with calculation of the water/total area ratio.

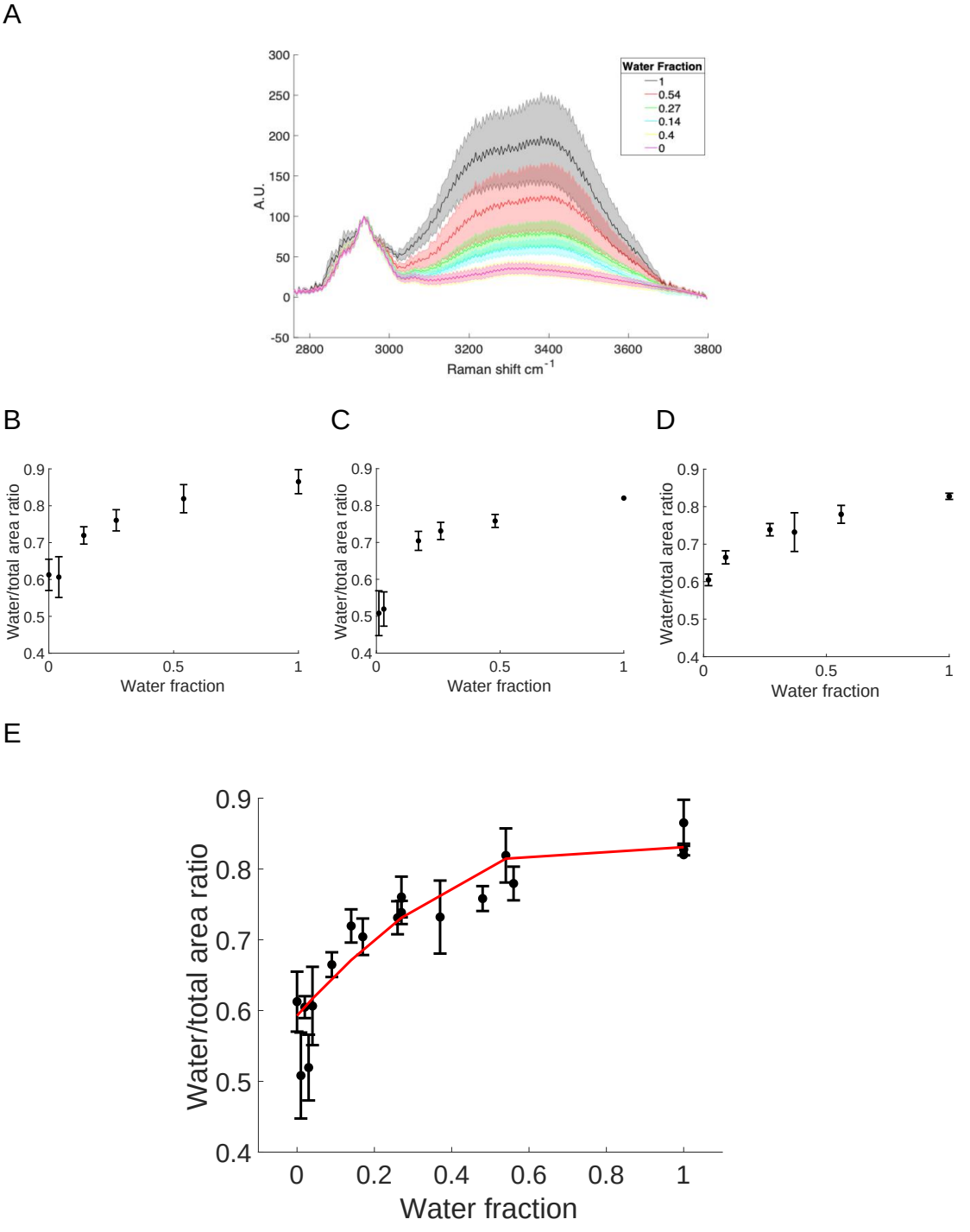


Figure 4-39 Results of serial dehydration of tumour specimens (n=3) and corresponding water/total area ratio

A. Spectral data of a single tumour specimen that underwent serial dehydration, with spectra for each corresponding water fraction plotted. Data line is mean of all spectra from the specimen at that water fraction (n=5 areas measured). Shaded area either side in the same colour is +/- 1 S.D. Spectra were baselined using 3rd order polynomial and normalised to 2935cm^{-1} peak. Graphs B (Specimen 1) C (Specimen 2) D (Specimen 3) and E (combined data of Specimen 1, 2 and 3) plots the mean water/total area ratio (represented by plotted points, error bars +/- S.D.) calculated from the spectra (n=5 measurements for each water fraction point, for each specimen) obtained at each water fraction plotted against the water fraction calculated by weight. Graph E has a second order polynomial line of best fit to demonstrate the 'dehydration curve'

Non-tumour specimen

Two non-tumour specimens underwent serial dehydration, however only one specimen was suitable for analysis due to an inability to accurately weigh one sample. The non-tumour specimen that underwent serial dehydration had a starting weight of 0.240 g and dry weight of 0.215 g, which was a 10% weight loss. This suggests that the, generally fatty, normal breast tissue had a low initial water content. The Raman spectra normalised to the protein/lipid peak demonstrated very little change in the water peak with serial dehydration, and analysis of the water/total area ratio suggested there was no trend in ratio with the water fraction (Figure 4-40).

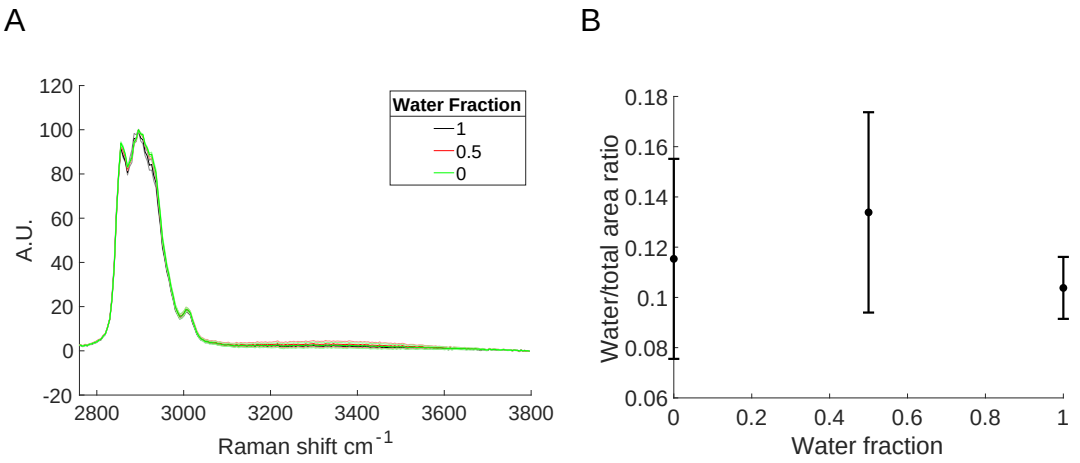


Figure 4-40 Results of serial dehydration of a non-tumour specimen with corresponding Raman spectra (n=1)

A. Spectral data of a single non-tumour specimen that underwent serial dehydration, with spectra for each corresponding water fraction plotted. Data line is mean of all spectra from the specimen at that water fraction (n=5 areas measured). Shaded area either side in the same colour is +/- 1 S.D. Spectra were baselined using 3rd order polynomial and normalised to 2935cm^{-1} peak. B. Scatter graph of the mean water/total area ratio (represented by plotted points, error bars +/- S.D.) calculated from the spectra (n=5 measurements for each water fraction point) obtained at each water fraction plotted against the water fraction calculated by weight.

Interpretation and implementation of results

The results from the tumour specimen dehydration curve demonstrates that HWN Raman, and using the water/total area ratio, can detect changes in water content, and it may be possible with the dehydration curve to quantify changes in water concentration in tumour specimens. The results from the non-tumour specimens suggests that HWN Raman cannot detect changes in water content in low water environments or highly lipid environments.

The dehydration curve from the tumour specimens could be used for interpretation of differences in water content between tumour specimens. As an example, there were differences in the mean water/total area ratio between lobular carcinoma (0.70) and mucinous carcinoma (0.87), looking at these values on the dehydration curve they equate to approximately a water fraction of 1.0 (for a water/total area ratio for mucinous carcinoma) and 0.25 (for a water/total area ratio for lobular carcinoma), suggesting there is a 75% difference in water content between mucinous and lobular carcinoma. This is in concordance with the understanding of the histopathological differences between the two types of tumour – mucinous tumours are mucin (a glycoprotein) producing and may be associated with a higher water content, and lobular carcinoma which infiltrate adjacent fat and so could be associated with a more fatty environment [357] which are associated with a lower water content.

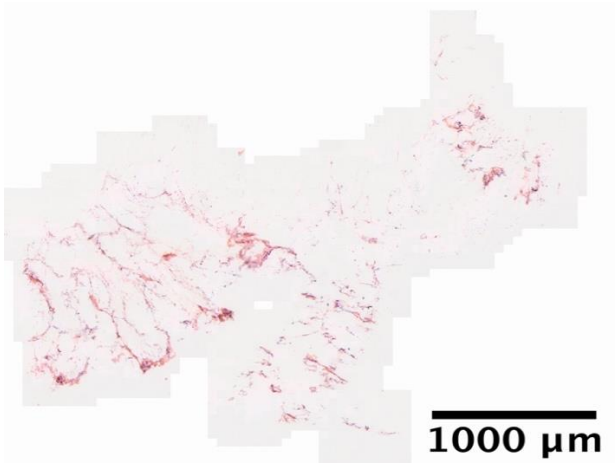
4.3.4 Raman micro spectrometry experiments

Non- tumour specimens

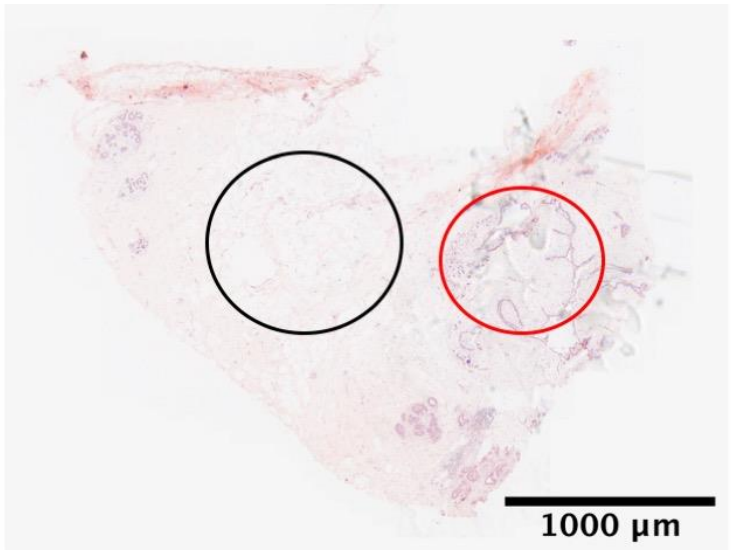
Review of H+E slides from non-tumour specimens demonstrated that a large number had areas of fat cells only, with very little histopathological features. Some non-tumour specimens however, while still predominantly fat cells, had areas of stromal tissue (Figure 4-41). It should be noted that the patients demonstrating these areas of heterogeneity were younger (age range 43-51) than the mean age of the sample population (67).

A non-tumour specimen had histopathological analysis followed by HWN Raman using Raman micro spectroscopy. The specimen was chosen for analysis because it was histologically heterogenous, and had good enough H+E staining to allow for accurate pathological analysis, the results are shown in (Figure 4-42).

A



B



C

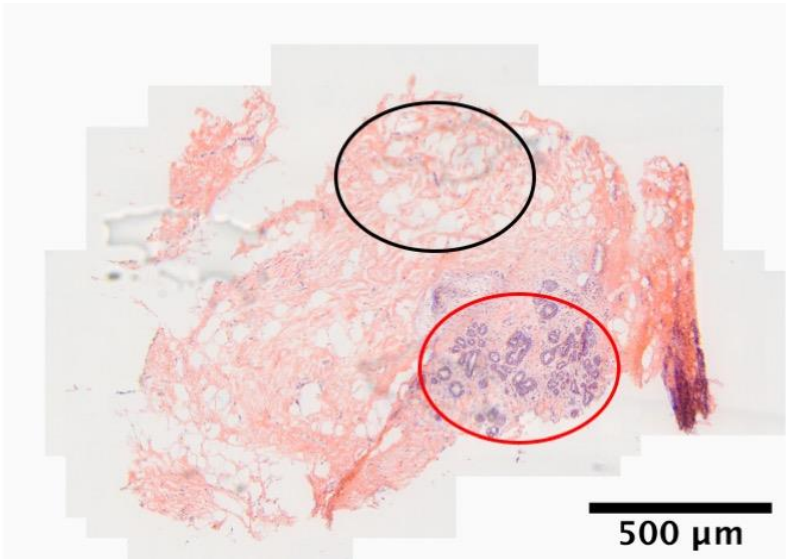


Figure 4-41 H+E slides of non-tumour specimens

A. Patient 14 (age – 51)) Specimen composed of mostly fatty tissue. Note that little of the H+E stain is taken up by the fat cells, and there are few histopathological features B. Patient 105 (age-43) and C. Patient 59 (age-44) demonstrate non-tumour specimens with a mixture of tissue types. Black circle denotes area of fat cells, Red circle denotes stromal tissue. Images taken under X 20 magnification

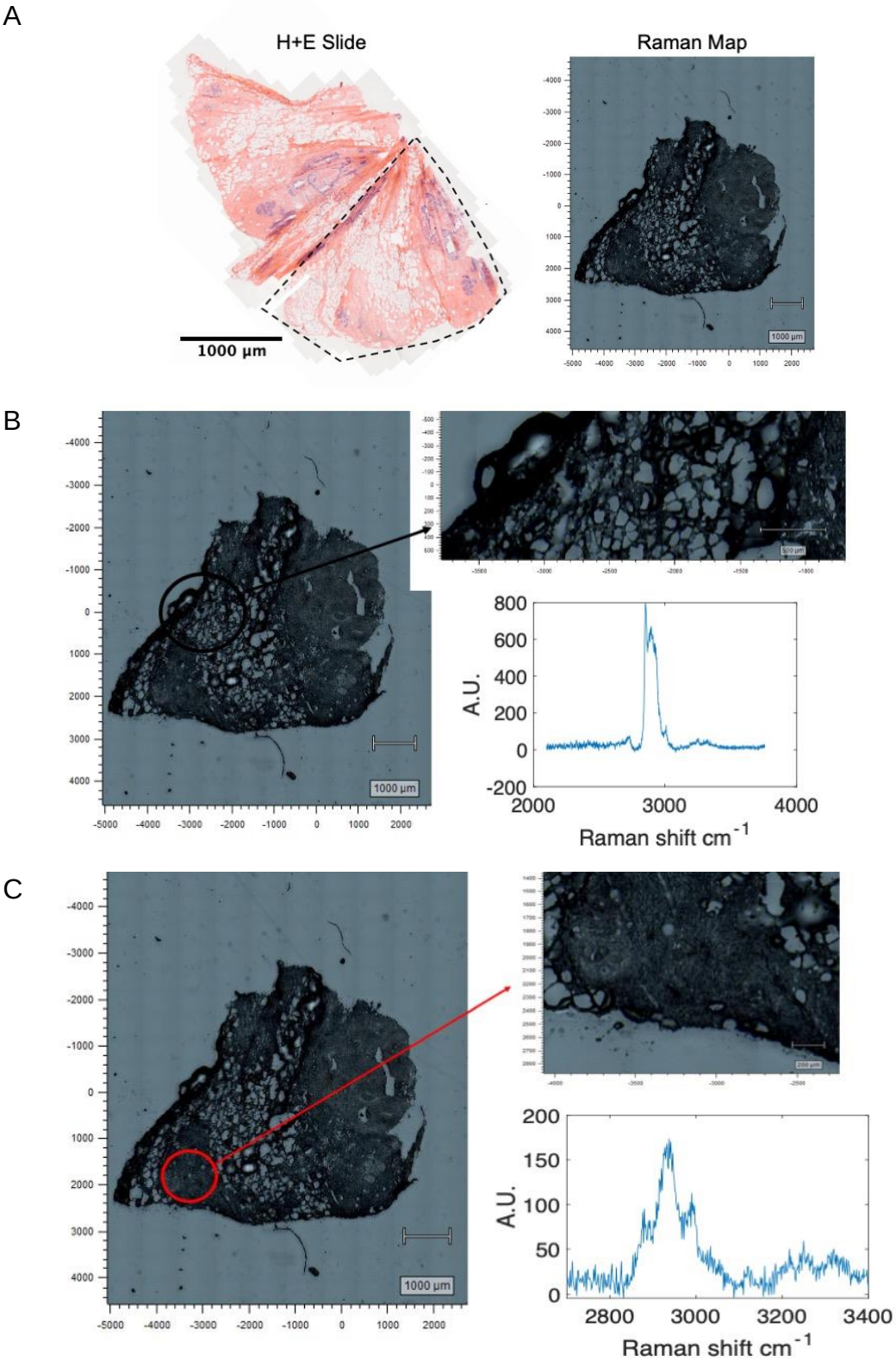


Figure 4-42 HWN Micro-Raman analysis of non-tumour specimen from patient 105

A. H+E slide of non-tumour specimen, black dotted line demonstrates the area of the corresponding Raman map. B. Area of fatty normal tissue highlighted in black at X5 magnification, and at X50 magnification with representative HWN Raman spectra showing the lipid peak. C. Area of stromal normal tissue highlighted in red at X5 magnification, and at X50 magnification with representative HWN Raman spectra showing the protein peak.

Tumour Specimens

Histopathological assessment of the tumour specimen indicated that the majority of tissue from tumour specimens was entirely tumour, examples are shown in Figure 4-43.

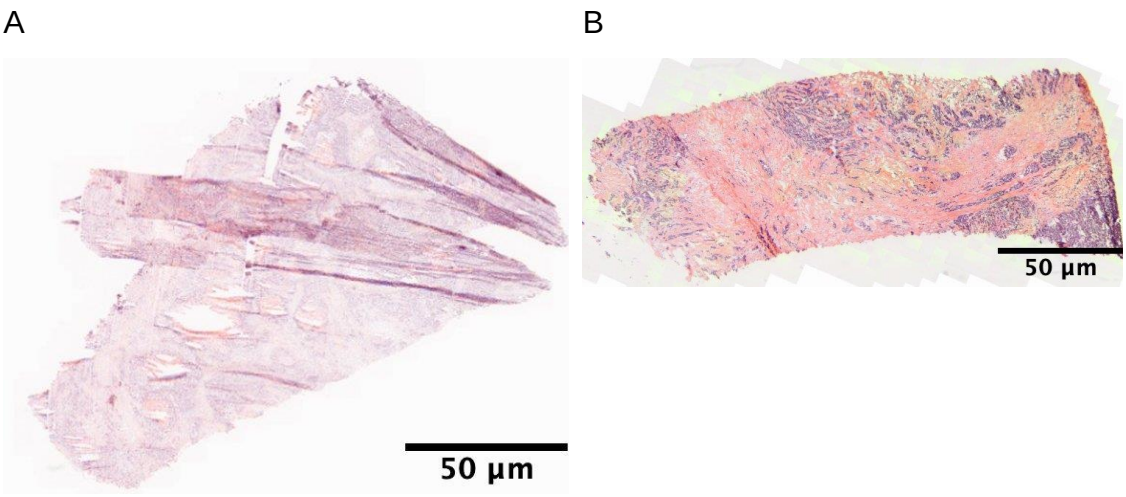


Figure 4-43 H+E slides of tumour specimens

A. (patient 17) Grade 3, Invasive ductal carcinoma, ER/HER2 -ve with no DCIS present. B. (patient 31) Grade 2, invasive ductal carcinoma, ER +ve/HER2-ve, associated with DCIS. Both slides demonstrate features of tumour throughout the entire histological slide.

However, there were some specimens that had areas of normal tissue within the specimens. Raman analysis of one of these specimens shows that the areas of the section that were identified as being tumour had a protein signal (matching the Raman signal seen from the tumour specimen fresh frozen sample measurements), and areas of the section that were identified as normal and non-cancerous had a lipid signal (matching the majority of the non-tumour specimen fresh frozen sample measurements) (Figure 4-44).

One of the specimens analysed (notably a lobular carcinoma), had histopathological evidence of tumour throughout the specimen, as is common in lobular carcinoma, there was cancer cell infiltration into surrounding fatty tissue[357]. Raman analysis of the specimen demonstrated areas of tumour that had Raman spectra with a protein signal, which would be expected in tumour according to our results from the fresh frozen specimens, but it also demonstrated areas of tumour with a strong lipid signal, which would likely be classified as being non-tumour based on the measurements gained from the fresh frozen specimens (Figure 4-45). This highlights that lobular carcinoma may be a potential source of diagnostic inaccuracy.

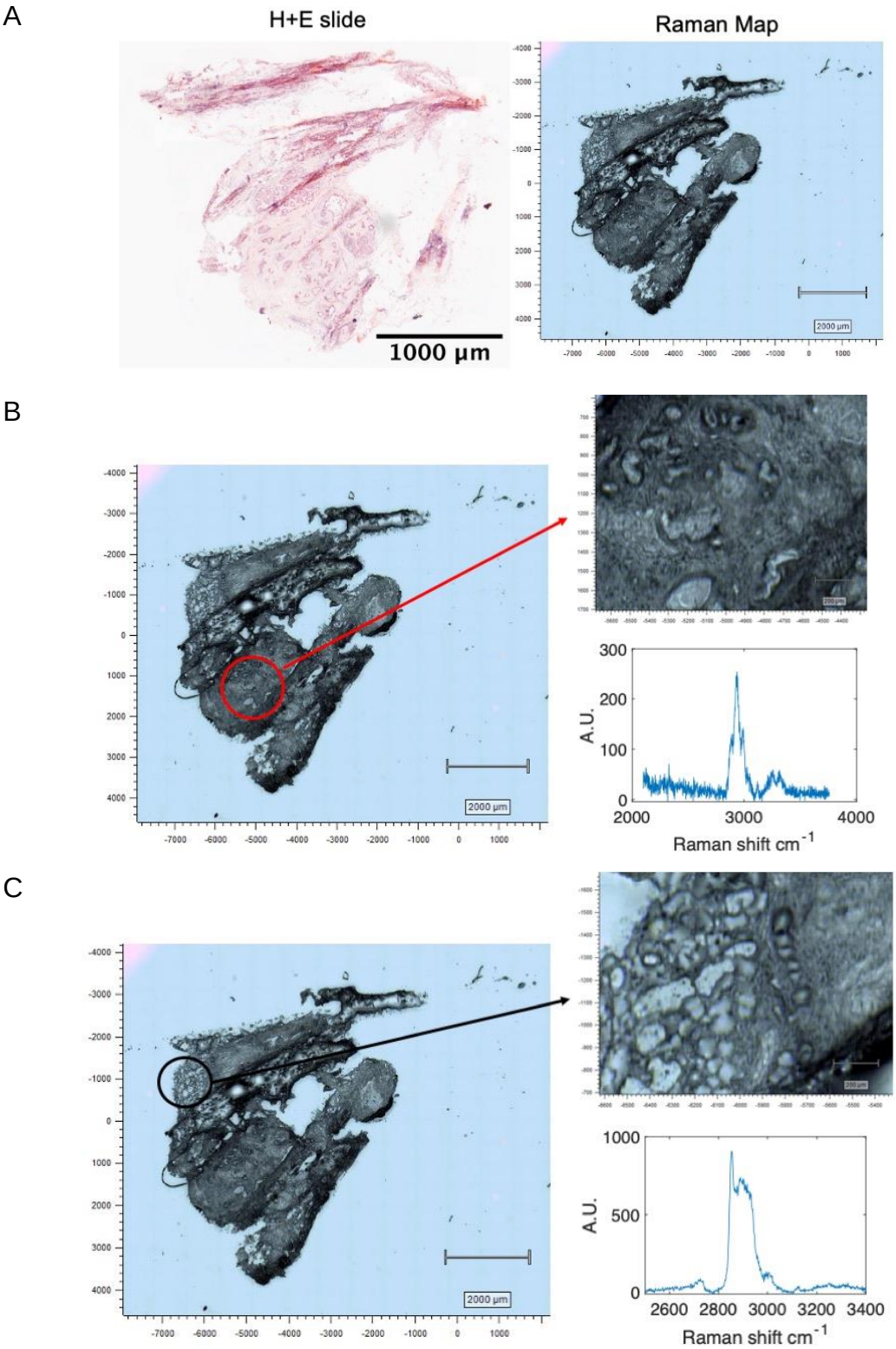


Figure 4-44 HWN Micro-Raman analysis of tumour specimen from patient 105

A. H+E slide of tumour specimen, a Grade 2 invasive ductal carcinoma, ER +ve, HER2+ve associated with DCIS, and corresponding Raman map. B. Area of tissue containing tumour highlighted in red at X5 magnification, and at X50 magnification with representative HWN Raman spectra showing the protein peak. C. Area of normal fatty tissue highlighted in black at X5 magnification, and at X50 magnification with representative HWN Raman spectra showing the lipid peak.

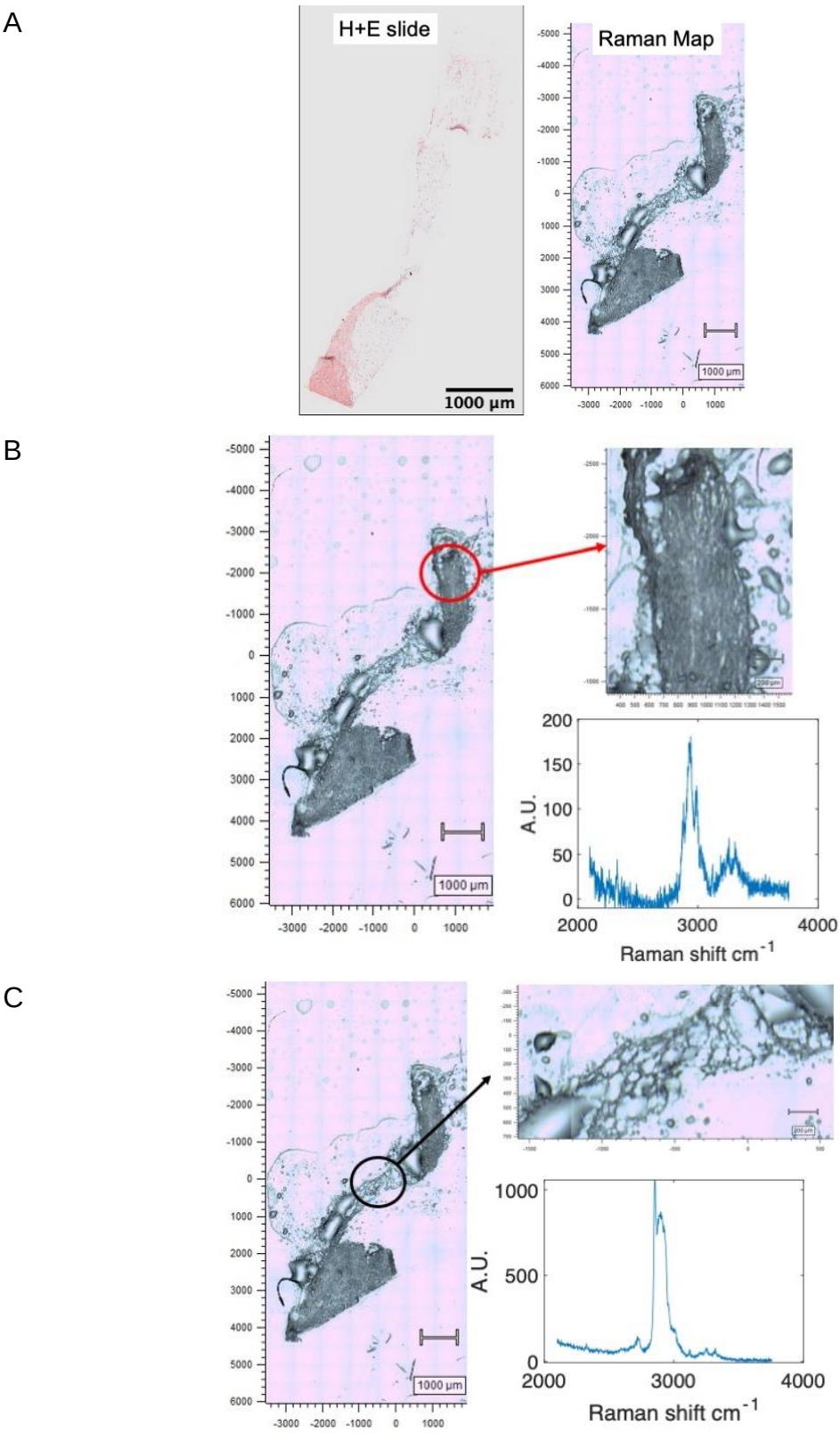


Figure 4-45 HWN Micro-Raman analysis of tumour specimen from patient 14

A. H+E slide of tumour specimen, a Grade 2 invasive lobular carcinoma, ER +ve, HER2-ve associated with DCIS, and corresponding Raman map. B. Area of tissue containing tumour highlighted in red at X5 magnification, and at X50 magnification with representative HWN Raman spectra showing the protein peak. C. Area of tumour containing tumour tissue highlighted in black at X5 magnification, and at X50 magnification with representative HWN Raman spectra showing the lipid peak.

4.4 Discussion

This chapter demonstrates the diagnostic ability of HWN Raman spectroscopy to differentiate tumour from non-tumour tissue in human breast tissue, which has been demonstrated for the first time. This is a promising technique that may be able to provide IMA in breast conserving surgery. The individual aspects of the results from the study will be discussed, and the clinical implications considered, with a view to what future work can be recommended to successfully translate the findings of this proof of concept study to a clinically robust method of IMA.

The results of this study have demonstrated that through analysis of the water/total area ratio as a measure of water content, and spectral analysis of both the fingerprint and HWN region a number of techniques can be utilised in the differentiation of human breast tissue with good diagnostic accuracy. The results suggest that the spectral data obtained using the NP3 system, a 785 nm excitation laser and InGaAs camera configuration, is unable to give specific pathological diagnoses, as there were no differences between pathological subtypes, however, tissue that had DCIS only was able to be differentiated from non-tumour tissue which is crucial for accurate IMA. The dehydration experiments have given an important insight into the role of measuring water content in breast tissue in the diagnosis of tumour tissue, and suggest that HWN Raman measures a change in the microenvironment of water more accurately than quantitatively assessing the change in water content. The biochemical changes seen from the Raman spectra to differentiate between tumour and non-tumour tissue using NP3 have been matched to the histopathological diagnosis using Raman micro spectrometry, which validates the conclusions regarding differences between tissue types.

4.4.1 Interpretation of results

4.4.1.1 Chemometric analysis

The results from this study highlight the biochemical differences between tissue types. Analysis of the differences in water/total area ratio between tumour (high water/total area ratio) compared to non-tumour tissue (low water/total area ratio) suggest that tumour contains a higher water content compared to non-tumour tissue. This difference in water content between tumour and non-tumour tissue has previously been suggested in human breast tissue with other optical

techniques, but it has not been used as the primary diagnostic marker [276, 297]. In this study we have quantified those changes in water content between the two tissue types and demonstrated that this simple marker is powerful enough to accurately differentiate between tumour and non-tumour tissue with an overall accuracy of >90% (depending on the exact data analysis protocol).

Peak assignation with values from the literature of the significant principal components identified that in both the fingerprint region and HWN region of the Raman spectrum non-tumour tissue had a high lipid profile and tumour tissue was predominately protein, and that there are the spectral components that differentiate between tumour and non-tumour tissue. Haka *et al.* used spectral data from the fingerprint region to differentiate between tumour and benign breast tissue, and noted that lipid and protein changes were central to the diagnostic algorithm that achieved good specificity of 93%, but a poor sensitivity of 83% due to misclassification of certain benign breast tissue types [186]. Previous data from the HWN region from breast tissue in rats [271] and human tissue [336] suggested these biochemical changes were also observed in the HWN region, but no studies have previously been performed on human breast tissue with a view to using the biochemical changes for diagnostic classification. Neither has a study concatenated data from both of the regions to produce a single spectrum from a single area that includes this data from both regions. The study outlined in this thesis confirms that both regions are evaluating similar biochemical changes, and the main spectral features in differentiating tumour from non-tumour tissue are protein and lipid changes (fingerprint region) and protein, lipid and water changes (HWN region).

The Raman mapping experiments validate the main chemometric findings from the NP3 system. The Raman mapping showed that areas of tumour, confirmed by histopathology diagnosis, were characterised by protein signal, and areas of non-tumour were characterised by lipid signal. This demonstrates that there is agreement in the chemometric data gained from HWN Raman spectra gained with the NP3 system with peak assignation from the literature and spectra from micro spectrometry that were matched to a formal histopathological diagnosis. This gives confidence to the assumption that ‘tumour’ signals were indeed from tumour tissue, and non-tumour signals were from non-tumour tissue.

A limitation of the Raman mapping method is that it used dried slices of tissue, and therefore no information on water content could be ascertained, and changes in water content could not be matched with the histopathology diagnosis. However, it was a consistent feature in the fresh frozen samples that protein rich tissue, which was matched with the histopathology diagnosis, had a higher water content.

4.4.1.2 The role of water content in differentiating tumour from non-tumour tissue

An original aim of this study was to investigate whether water content was a marker of differentiating tumour from non-tumour, and whether this could be quantified to assess if changes in tissue water content could be used in the assessment of pathological margins.

The results from this study suggest that differences in water content are able to differentiate tumour from non-tumour. The spectral data from the HWN region, particularly when normalised to the protein peak demonstrate that tumour specimens have a much higher water peak than non-tumour specimens, and when using the water/total area ratio as a measure of water content, tumour specimens have a significantly higher water content than non-tumour specimens. Using the water/total area ratio alone could accurately differentiate between tumour and non – tumour specimens, with an accuracy of 89-91%, confirming that the difference in water content between tumour and non-tumour tissue is substantial, measurable, and diagnostic.

The dehydration experiments had some success in quantifying changes in water content. Using tumour tissue, which had a high initial water content, the water content of the tissues was sequentially measured, and the associated Raman signals obtained, from which a measure of water content at that time (the water/total area ratio) could be calculated. This produced a calibration curve allowing HWN Raman spectra to be taken of a breast tumour specimen, the water/total area ratio to be calculated, and the calibration curve used to assess the water content of that breast tumour specimen, the example of a lobular carcinoma versus a mucinous carcinoma was used. This demonstrates that

quantification of changes in water content can be performed in tumour tissue. It should also be noted that it is only relative changes in water content that can be measured, as the absolute starting water concentration of the tumour specimen was not calculated, and, as stated in chapter 3, other absolute measures of water concentration may not be relevant to the microenvironment of breast tissue.
Measuring the relative differences in water can be used to improve HWN Raman in providing IMA, as it has been utilised in skin cancer [295], and oral SCC [174]. It also allows correlation with HWN Raman findings with other modes of investigation and diagnosis that use changes in water content, such as DOS [276], Terahertz imaging [288] and MRI [358], and so allow the findings with these modalities be translatable to Raman research, and vice versa.

The dehydration experiment results also demonstrated what was predicted in chapter 3, that the water/total area ratio relationship with water content is different according to the tissue microenvironment. This means that if the same water/total area ratio is measured in an area of tumour tissue and non- tumour tissue the two areas will not have the same water concentration. An obvious illustration of this in the dehydration experiment results is that when the specimen is in a state of complete dehydration (where the water fraction =0), for non-tumour tissue, the water/total area ratio is 0.12, and for tumour tissue it was between 0.5 and 0.6. This is the only point at which the two types of tissue can be directly comparable as having no water content, whereas at the other points of measurement they will have varied according to the initial water concentration. One of the reasons for this particular finding is that protein has a peak at 3280 cm^{-1} due to NH stretch [345], which is present in dehydrated protein tissue, but not in lipid tissue, which will contribute to the 'water' part of the water/total area ratio giving protein a higher reading at this complete dehydration compared to lipid tissue. However, this is not the only reason for the findings, as lipid is more Raman active than protein, and lipid gives a broader Raman peak, both of which would also alter the relationship with the water/total area ratio, even if the water peak was identical in both tissue types. This confirms what has been discussed in Chapter 3, that the water/total area ratio, and the change in HWN spectrum according to changes in water content, is dependent on the microenvironment of the water, and so changes between that of a predominantly protein environment (tumour) and lipid

(non-tumour) environment, meaning the same dehydration curve cannot be used for calibration for both environments.

There are limitations to the methods of the dehydration experiments, which need to be considered. The experiments were done in fresh frozen samples, which may have a different initial water content to fresh breast tissue, and the thawing process could alter the water content rapidly. The distribution of water and type of water (bound versus unbound) could also be different between fresh frozen tissue and fresh tissue – chapter 3 demonstrated that this has an effect on the HWN Raman spectrum, and subsequently the water/total area ratio. The dehydration curves also demonstrate that the majority of the dehydration (as measured by loss of mass) occurred between the initial and second measurement (over the course of the first hour), with some specimens losing almost 50% of their mass during this period. The specimens were so small the rate of dehydration was rapid. However, despite these limitations, the findings are consistent and repeatable (in the tumour specimens, 3 different specimens gave remarkably similar results). The findings would ideally be validated by repeating these experiments in larger pieces of fresh breast tissue that could be measured more frequently. This would be logistically very difficult.

Comparing the dehydration curve of the breast tumour tissue and the dehydration curve of the pork meat tissue, and the breast non-tumour tissue and the pork fat tissue in chapter 3, the curves are very similar in shape, rate of dehydration, and findings. The pork tissue did not have any of these limitations and gave similar results, suggesting that the findings in the fresh frozen breast tissue are valid and may not be altered if performed in more ideal tissue.

These findings demonstrate that HWN RS differentiates between the tumour/ non-tumour specimens as an analysis of the microenvironment (whether it is a protein or lipid environment) and the effect of that microenvironment on the water peak, rather than by quantifying the changes in water content. Therefore, when measuring different tissue types containing protein and/or lipid, the protein and lipid peak should be considered along with the water peak for diagnosis, rather than just using the water peak.

Within the tumour (protein) environment, there can be differentiation and quantification in changes in water content as the water/total area ratio changes in response to changes in the water content. These more subtle changes in the tumour environment may assist in the diagnosis of pathological margins. Previous work has suggested that in breast tissue, and breast cancer, rather than there being an abrupt change in the water content, there is a steady change in the water content when measuring from a higher water environment (tumour) across the border of the tumour to low water environment (non-tumour) [293]. This theory has been examined with HWN Raman spectroscopy in the diagnosis of oral SCC margins [228], but has not been investigated for the diagnosis of breast pathological margins. In these previous studies, the effect that the change in protein and lipid microenvironment going from tumour to a non-tumour environment would have on the ability to interpret the difference in water peak was not explored – and our results suggest this is crucial to understanding the differences in water content in different microenvironments. The results from our study suggest that HWN Raman spectroscopy is able to detect and quantify changes in water content, allowing the theory of water changes at the margins of breast tumour tissue to be investigated in the future.

4.4.1.3 DCIS only tissue

The ability to differentiate between DCIS only tissue and non-tumour tissue is paramount for IMA. Despite this, in a number of studies intending to provide IMA, DCIS only tissue is often excluded from analysis, which results in a major limitation of the clinical relevance of the study [156, 180, 185]. The majority of positive margins are due to DCIS only [187], and not invasive tumour, and so the NP3 system must have the ability to identify DCIS only tissue as requiring resection if it is to be clinically useful.

The initial analysis on all tumour tissue (which was composed of mainly invasive tumour tissue, but 3 DCIS only specimens) suggested that this was possible. The majority of tumour tissue (77%) was associated with DCIS, and it is highly probable that when measuring the tumour specimens, the associated DCIS was being measured as well, and contributing to the tumour signals. When the detailed histopathological analysis was performed for the Raman micro spectrometry experiments, DCIS was often adjacent or within areas of invasive

tumour. There was also no difference identified between water/tumour area ratio, nor on principal component spectral analysis when pathological subgroup analysis was performed between those associated with DCIS, and those that were not.

In addition to this a subgroup analysis was performed on the specimens from 3 patients that had DCIS only (and no invasive element) in their tumour specimen. This analysis demonstrated that there were obvious visual differences between the DCIS only tissue and non-tumour tissue. The biochemical differences were the same as the analysis between the all tumour data (of invasive tumours and DCIS only) and non-tumour specimens, DCIS only tissue was predominately protein with a high water content, compared to the non-tumour tissue of lipid with a low water content. Analysis of these spectra using the water/total area ratio and concatenated spectral data suggested the potential for satisfactory diagnostic ability to differentiate between DCIS only tissue and non- tumour tissue. A limitation of this subgroup analysis is the small number of patients and therefore the statistically significant findings should be interpreted with caution. However, these results do suggest that the basic chemometric principles that are used to differentiate tumour from non-tumour specimens hold true for differentiating DCIS only and non-tumour specimens. Although it would require a larger number of DCIS only tissue to be conclusive, these results suggest that NP3 has the ability to differentiate between DCIS only and non-tumour tissue, a crucial attribute for IMA.

4.4.1.4 Comparing data analysis techniques

Based on the chemometric findings, a number of data analysis techniques were used to assess the ability to differentiate between tumour and non-tumour tissue based on the Raman spectra. Each technique had broadly similar results, with generally good overall accuracies between 89-93%, and a broader range of sensitivities 87-96% and specificities 86-96%. An important feature of the analysis is that all methods underwent cross validation, with testing of the prediction accuracy with a ‘prediction’ set of data that was separate to the data used for ‘training’ the diagnostic model. This suggests that if the diagnostic model was used in an entirely new set of spectral data, the results would be similar.

A benefit of having performed a number of methods of spectral analysis is that the protocol that would be suitable for clinical use can be assessed. The technique for use in a clinical environment for IMA would take into account what is needed from the test in order to get the desired clinical outcome. For example, a test that has very high sensitivity would reduce the re-excision rate considerably, but if there was a low specificity of the test it may result in an unnecessary removal of breast tissue, which is associated with poor cosmesis. Equally a highly specific test would guard against the unnecessary removal of breast tissue but may not be sensitive enough to reduce the number of re-excision operations significantly. Therefore, each analysis technique should be considered in turn and assessed as to how it could be used for IMA.

Water/total area ratio

The water/total area ratio is a simple technique for assessing differences between spectra. It requires minimal pre-processing (fluorescence baseline subtraction only), and a simple calculation of the AUC of two peaks (water and the total HWN spectral region) in order to calculate the ratio. This ratio is also a single figure that can be used to characterise a spectrum. Using simple data analysis such as this may have real strengths in the clinical environment for IMA, as it would require less processing power, be quicker than more complex techniques and be easier to devise diagnostic algorithms.

A use of the water/total area ratio is by using different voting thresholds to provide a single dichotomous answer for each specimen. In a clinical setting this method could be used to take a number of measurements for the entire resected edge of a specimen and using the data to provide an outcome of whether that resected edge had a positive margin and further resection was required or not. The results demonstrated that it was possible to use a defined cut off of water/total area ratio to classify a spectrum as tumour or non-tumour, the proportion of the number of ‘tumour’ readings could then be taken into account to ‘vote’ if the specimen (or resected edge) was positive. The cut-off could be varied according to the clinical outcome required – and whether greater or lesser sensitivity or specificity was required needs more research. The two voting methods give very different clinical outcomes (Voting threshold 1 sensitivity - 87.5%, overall accuracy 91.7%, Voting threshold 2 – sensitivity 95.98%, overall accuracy 88.0%). The limitations are

evident, voting threshold 1 (where the majority spectrum classifies the entire specimen) could lead to small areas of tumour being 'missed', and voting threshold 2 (where a single tumour spectrum classifies the entire specimen as tumour) could lead to an unnecessary amount of tissue being excised. However, these methods represent a clinically relevant method for data analysis and giving a clear outcome for a resected edge, which is essential for IMA.

Using the W/TAR difference is a different way of utilising the ratios to take into account variations between patients physiological 'normal'. The analysis of the normal breast tissue demonstrated significant variations in the Raman spectra between patient specimens, particularly with age. Using a 'calibration' reading from obviously normal tissue, and comparing the possible tumour readings to this 'calibration', there is the possibility of increasing the accuracy of the device. The results in this study are very dependent on the 'cut off' used to define a tumour or non-tumour, though this could be adjusted for the desired objective of analysis. Though, interestingly, whatever the cut off used, the overall accuracy of this technique was equivalent to the other methods that did not take into account the individual patients baseline normal tissue, perhaps suggesting that this does not affect the diagnostic accuracy significantly. This method of 'calibration' of the machine prior to taking margin readings is already performed in another device that is used for IMA – the ClearEdge [266], suggesting others have thought it important to take into account the individual baseline of the patient. However, ultimately the results from the ClearEdge were disappointing, and in a number of cases there was surgeon error in using the device to achieve the 'calibration' reading. It may be that a calibration reading step could over-complicate the measurement protocol, without giving sufficient additional diagnostic accuracy.

There are limitations to using the water/total area ratio. One, explored above, is that the physiological normal, non-tumour, readings can overlap with the ratio of tumour tissue. Particularly younger, pre-menopausal women may have higher water/total area ratios in normal tissue and have a greater incidence of false positive readings using this technique. Another consideration is that although in this analysis the water/total area ratio diagnostic accuracy was similar to the more complex spectral analysis, it may be that there are subtle changes in the spectral analysis that would be missed if only the water/total area ratio was used for

diagnosis. Spectral analysis gives greater tissue biochemical information and it may be that minor variations did not reach statistical significance in this setting due to the relatively small number of spectra, but with a greater sample number these minor variations may become significant and provide a greater diagnostic accuracy.

Spectral analysis

Spectral analysis was performed on the fingerprint region, HWN region and the two regions concatenated. The diagnostic performance using the mean spectra from the fingerprint (accuracy 92.7%) was slightly reduced compared to the HWN (accuracy 93.2%) region. Interestingly the accuracy from the concatenated spectra was the same as that of the FP only region (92.7%), suggesting that the addition of the fingerprint region reduces accuracy compared to the diagnostic information of the HWN region alone.

The chemometric analysis suggest that similar biochemical differences in the tissue are used to differentiate between tissue types, but that the HWN region provides information on water content in addition to the protein/lipid changes seen in the fingerprint region, suggesting that the change in water content information may be responsible for the increased diagnostic accuracy when using the HWN region.

When using the spectral data for differentiating tissue types, the concatenated spectra could be used. There are no negative aspects of using the entire spectrum that is available when measuring using the NP3 system, and both fingerprint and HWN regions are captured using the NP3 system without having to change any hardware such as laser excitation wavelength or filters – unlike other systems that have used both regions for tissue diagnosis [299, 304]. This means to capture both regions does not take any more time or expense compared to capturing just a single region. Although, from this study, there appears to be no diagnostic benefit to using both regions, it may be that further work with a greater number of samples exposes subtle spectral changes that allow an increased diagnostic accuracy while using both regions for data analysis.

A limitation of using the entire spectrum and performing spectral analysis is that, although in this study the acquisition time allowed for obtaining both regions of the spectrum with good spectral resolution, if acquisition time was decreased significantly (to allow more rapid IMA) it may mean the entire spectral region was not adequate for analysis. In the spectral analysis a sequence of principal component analysis (PCA) was used to calculate scores which were fed into a linear discriminate analysis (LDA). This is a complex method of analysis that would take longer than for calculation of the water/total area ratio and may require greater computing hardware. A way to decrease this burden of spectral analysis would be to reduce the number of principal components used. In all cases it was the first 2 principal components (with a minor contribution from PC3) that accounted for >95% of variance in spectra contributing to tissue diagnosis. The spectral analysis could be reduced to the first 2 or 3 principal components to reduce the data processing time, however, it may come with a decrease in diagnostic accuracy.

Another limitation of spectral analysis is that it is reliant purely on mathematical modelling to decide the best way to analyse the tissue specimens, without the ability to change it according to clinical need. Using the water/tumour area ratio, the threshold for diagnosing a spectrum as ‘tumour’ is easily manipulated to achieve the desired clinical aim, whereas using a PCA fed LDA model it is not so easy to ‘fine-tune’ the model to the desired clinical outcome.

The use of mean data

A consistent finding in all the spectral analysis, regardless of method, was that the mean data (that is obtaining a mean spectrum or ratio for each specimen, so each specimen had a single output) gave an improved diagnostic accuracy compared to analysing each individual spectrum as a discrete reading. The most apparent example is for spectral analysis of the HWN region, where using individual spectra gave an overall accuracy of 90.3%, and using the mean spectra gave an accuracy of 93.2%. The reason for this is likely to be when analysing the data as single spectra, if a large number of specimens each had a small number of ‘discordant’ spectra (i.e. non-tumour measurements in a tumour sample), this would give a large number of ‘discordant’ spectra overall that would be classified incorrectly. If taking the mean of each specimen, these small number of

'discordant' spectra would be averaged out, compared to the large number of e.g. tumour spectra in tumour specimens. The Raman mapping experiments demonstrated that within 'tumour' specimens, there were areas of non-tumour, so this may explain why there were a small number of 'discordant' spectra which negatively affected the diagnostic accuracy when assessing the spectra individually.

Analysing the data from these experiments using the mean data has greater statistical validity, as for each patient specimen there is a single outcome. When analysing each spectrum individually, in the diagnostic models a number of measurements are originating from the same tissue, and if, theoretically one patient contributed 100 measurements, and another only 10 (which did not occur, the variation in the number of contributed measurements was +/-2) that would unbalance the model in favour of the type of measurements from the patient contributing 100. However, from a practical point of view of providing IMA, it is not necessarily clinically relevant to use a 'mean' spectrum. The mean could be the mean of a number of readings from a small area, or the mean of all readings from an entire resected surface. The area over which to average the readings needs to be determined, and would be reliant on the sampling area and properties of the device that would be performing IMA. The findings from chapter 3, and observing the raw 'counts' from spectra gained from this study, is that the protein signal is weak and lipid has a relatively strong Raman cross section, and so small areas of tumour (protein rich) could easily go undetected with this method of data analysis. The main danger of using a mean spectrum over a large area, such as the entire resected surface, would be a reduced sensitivity that small areas of abnormal tissue would be averaged out and not be detected. It may be that providing a mean spectrum over a small area would improve accuracy, account for minor physiological variations, and improve the diagnostic accuracy of the analysis.

Suggested protocol for data analysis

It is important that the data analysis of Raman spectra can be processed and utilised to provide a clinical outcome, which is easily interpreted by a surgeon, and informs whether a resected edge is 'positive' (i.e. tumour identified) and

further resection is required, or a resected edge is ‘negative’ (i.e. no tumour identified) and no further resection is required.

Based on the results and subsequent discussion of the various data analysis techniques of the spectral data obtained with NP3, the following sequence is suggested as the most accurate, and clinically relevant, method of data analysis (Figure 4-46):

1. The resected edge of the specimen has all Raman measurements taken.
    - The measurements are obtained in a systematic order, with the resected edge divided into ‘sections’ that allow grouping of close by measurements together. The ability to do this would depend on the final design/ measurement protocol of the probe to be used.
2. Pre-process data
    - Fluorescence background subtraction using 3rd order polynomial for HWN region, and 6th order polynomial for FP region
3. Calculate the water/total area ratio from the HWN region of each spectrum
4. **Primary classification** - the mean water/total area ratio for each section is used to classify the section as tumour or non-tumour with a low threshold of water/total area ratio (e.g. 0.4 - to provide high sensitivity).
5. **Secondary classification-** for all sections classified as tumour, analyse the spectral data to provide high specificity.
    - All the spectra from a section identified as ‘tumour’ by the first classification would be used in the reclassification of that section
    - The concatenated spectral data has min/max normalisation followed by PCA fed LDA analysis to re-classify if the section is tumour/ non – tumour.
6. **Final classification**- All section outcomes of resected edge are processed with a voting threshold analysis, if there is one or more area classified as ‘tumour’, the entire resected edge is classified as ‘positive’ and further resection is required.

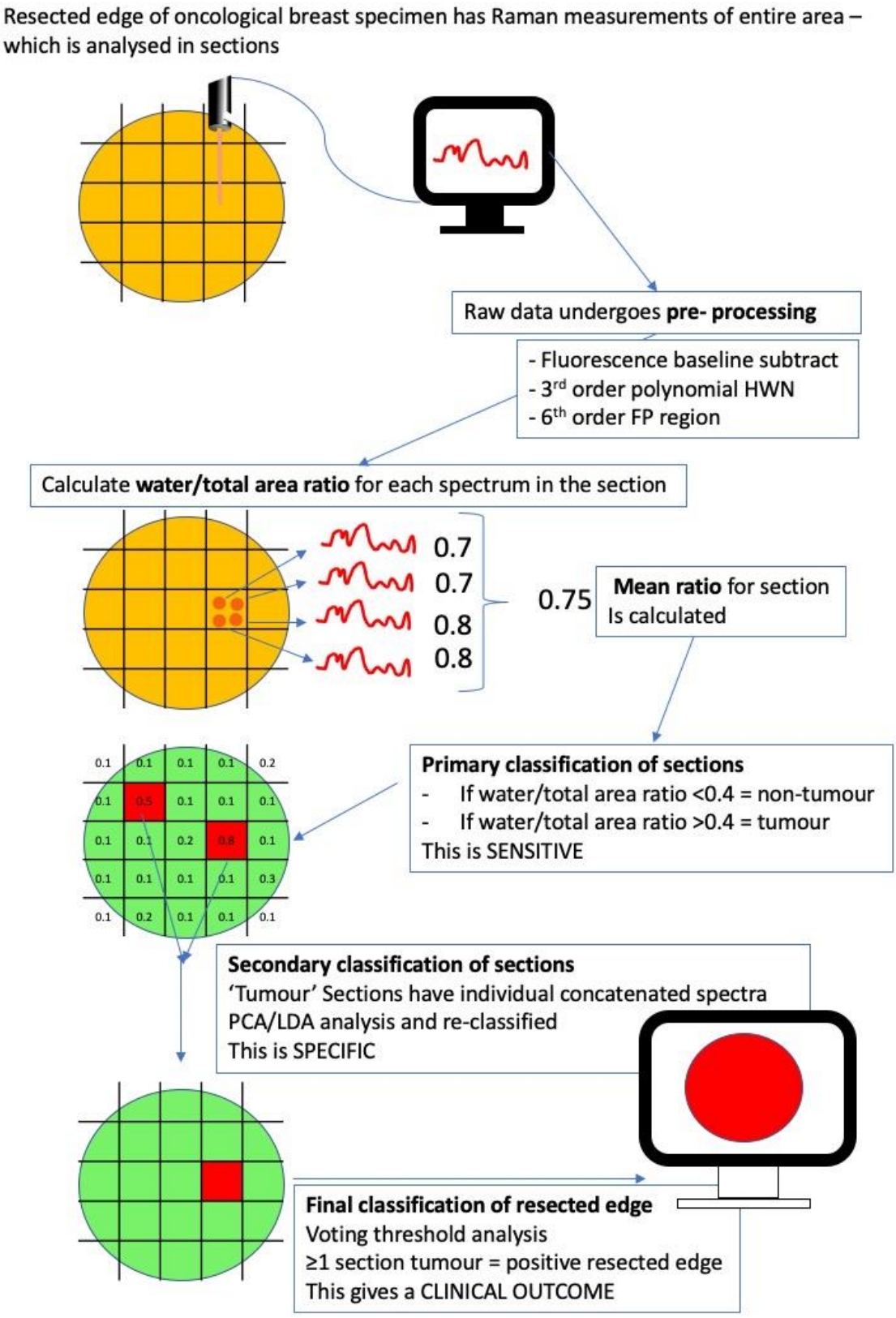


Figure 4-46 A schematic representation of how the suggested data processing algorithm could be used on a specimen resected to give a clinical outcome

4.4.2 Limitations of methods

4.4.2.1 Demographics of study patient and tumour characteristics

The demographics of the patients were broadly representative of patients undergoing breast cancer surgery in normal clinical practice. The age of patients (mean age 67) is representative of the age of patients diagnosed with breast cancer seen in clinical practice, the majority being post- menopausal [359].

A vital inclusion criteria in the study – that is tumours had to be greater than 2 cm to allow core biopsy and tissue sampling- leads to selection bias in the tumour characteristics of the specimens measured. 2 cm tumours are usually symptomatic- 89% of patients in the study group were diagnosed via a symptomatic route of referral; whereas only 75% of all breast cancers present with symptoms [41]. 2 cm tumours are also more likely to be removed by mastectomy rather than breast conserving surgery – 48% of the study group underwent mastectomy as opposed to breast conserving surgery. Mastectomy rates vary, but account for around 30-40% of breast cancer operations [6]. This means that larger tumours that would undergo a procedure not requiring IMA are slightly over-represented in this population study, when compared to the population that are likely to require IMA, that is, more screen-detected, smaller tumours undergoing breast conserving surgery. The result of this is that the tumour characteristics of the study population are possibly not representative of the population likely to require IMA. Screen detected tumours are more likely to be smaller, and diagnosed and treated at an earlier stage [359]. The screen detected population also present the majority of DCIS only disease which is often treated with breast conserving surgery, and represent a large number of patients that require re-excision of margins [187]. Therefore, early-stage disease and DCIS only disease is under-represented in the population study. The under-representation of DCIS only disease needs to be considered in future work. The initial assessment suggests that the NP3 system is capable of differentiating between DCIS only tissue and non – tumour tissue, but the number of such patients was small (n=3) and so further work does need to be performed to confirm this.

Despite these limitations the number treated by breast conserving surgery are only marginally below what would be expected. There is also a broad range of

breast cancer disease, and the carcinoma type (with the majority being Ductal), carcinoma grade (majority Grade 2) the receptor status (with the majority being ER+ and minority HER2+ve) is broadly in keeping with the tumour characteristics of a typical patient undergoing breast conserving surgery. Therefore, the study population is similar enough that it is probably representative, and so the data drawn from these studies is likely to be translatable to a general breast cancer population undergoing breast conserving surgery.

4.4.2.2 Specimen type

The specimens that underwent measurement had been taken from areas that were macroscopically tumour or non- tumour in the view of an experience healthcare professional (pathologist or associated practitioner). The review of the histopathology H+E slides suggest that the tissue samples were mostly composed of homogenous tissue types that had been correctly categorised as tumour or non- tumour. The benefit of using these samples is that the study specimens are similar to the type of tissue that is measured in IMA. They are larger pieces of tissue, over which a number of measurements were obtained, and a diagnosis made based on the Raman spectrum – similar to IMA.

The limitations of the tissue type are that they were fresh frozen, then thawed prior to measurement. This is unlikely to have affected the composition of the specimen with regards to protein or lipid but is likely to have affected the water content. Indeed, during the dehydration experiments the specimens lost almost 50% of their mass (in presumed water dehydration) in only an hour. However, the measurements took less than 5 minutes to complete, so significant amounts of dehydration may not have occurred. This is evidenced by the fact that the differences in water content between tumour and non-tumour tissue could still be measured, were significant and used for diagnostic purposes as demonstrated by the diagnostic accuracy of using the water/total area ratio. It is also true that in fresh breast tissue, when performing IMA, there would be dehydration from the resected specimen edge in the clinical environment. Therefore, if the water content remains a major spectral feature to differentiate tumour from non-tumour even in tissue that has been freeze/thawed, is small, and has a rapid dehydration time, may suggest that these are robust findings that would be valid in fresh breast tissue. The effect the freeze/thaw cycle had on the water content in the

specimens is unknown, and measurements in fresh tissue are necessary to understand if the findings regarding water content are valid in fresh breast tissue, which is the tissue type IMA would take place in.

A further limitation of the sample types is that they were all homogenous tissue and the tumour sample was taken from the centre of the tumour. This means the periphery of the tumour – which is where a positive margin would most likely be – was not necessarily included in the sample. Previous studies have investigated the 'borderline' regions of tumours, describing a gradual decrease in water content the further from the tumour is measured [228, 293]. It may be this gradual decrease in water content that could be measured to provide IMA. This study has demonstrated the basic science underpinning this technique – that tumours have higher water content, this can be measured using HWN Raman, and changes in tumour water content quantified using the water/total area ratio. Further work with access to take Raman measurements from the tumour and the borderline region into non-tumour tissue is necessary to see if this can provide IMA.

The non-tumour specimens did not undergo histopathological analysis and were all considered to be a homogenous group of 'normal' tissue. On review of a select few H+E slides of the non-tumour tissue, they were confirmed to be normal with no particular histopathological features. 'Normal' breast tissue is a broad variety of tissue encompassing a number of pathological features, from fatty tissue which has few histological features, to the myriad of benign breast tissue changes named under the umbrella term 'Aberrations in the normal Development and Involution of the Breast' (ANDI) [360]. These all have their own specific histopathological features which may have particular Raman signals associated with them. A previous study using fingerprint Raman demonstrated that these benign abnormalities could affect the diagnostic accuracy of the technique, as there was overlap in the Raman spectrum of these benign findings and tumour, particularly in fibrocystic change and fibroadenomas [262]. However, it is not known if any of the non-tumour tissue in this study contained these histopathological variations, therefore, the diagnostic algorithms cannot take these into account.

4.4.2.3 Methods of Analysis

Using the NP3 system for this study demonstrated the potential for performing IMA. Each measurement was simple to perform and would not require a specialist operator to perform, requiring only light contact between the needle probe and the sample. If necessary, the needle could have a single use covering over it to allow the actual needle to remain clean between patients.

There are limitations to the method of analysis. The NP3 probe had sufficient sampling volume (5.19×10^4 μm^2) to measure small biopsies of tissue, but to sample an entire resected edge of a breast tissue specimen would not be possible. The time taken for each measurement was 25 seconds (5 second acquisition with 5 accumulations), to measure such a small area this is not a clinically relevant period of measurement. The reason for this was to gain accurate diagnostic information, as this was a proof of principle study. The results suggest that the HWN region alone gives sufficient diagnostic information, and the water/total area ratio is as accurate as full spectral analysis. The water/total area ratio could still be calculated with a reduced spectral resolution, and so the acquisition times could be reduced significantly, while maintaining diagnostic accuracy to make the method more clinically relevant.

Although the NP3 system gave more clinically relevant readings than the micro spectrometry analysis, the readings obtained with the NP3 system could not be directly correlated with the histopathological features. The micro spectrometry analysis of the tissue sections demonstrated that there are some ‘normal’ areas of breast tissue within the breast tumour specimens. When measuring with the NP3 it was not possible to know if a ‘tumour’ area or ‘normal’ area was being measured from the tumour specimen, so ‘normal’ readings may have been obtained from a tumour specimen, which would have been classified as ‘false negative’ in the prediction models (when the readings were actually true negatives), and so the diagnostic accuracy may be an underestimate.

4.4.3 Tissue types presenting a diagnostic challenge

4.4.3.1 Pathological subtypes

The analysis of tumour specimens investigating spectral differences between pathological subtypes demonstrated that there were no significant differences between subtypes. This means the NP3 system cannot differentiate between different tumour characteristics. This is not a limitation for IMA, as the differentiation only needs to be tumour versus non-tumour, but if the probe were to be developed for other clinical needs such as needle biopsy, it may be relevant that it is unlikely the NP3 system could provide a specific diagnosis.

4.4.3.2 'Normal' tissue

Breast tissue composition varies between patients according to a number of physiological factors including age [283], menopausal status [282] and hormonal factors such as menstrual cycle and exogenous hormones [283, 285]. The demographic data that had been collected on these samples allowed analysis of the differences in non-tumour tissue according to age and menopausal status. These are independently associated with changes in breast tissue composition[284], however there is obviously a link between these two characteristics, as post-menopausal patients are generally older.

The analysis found significant spectral differences between the non-tumour specimens of younger and older patients and pre-and post -menopausal patients. Younger patients, and pre-menopausal patients had significantly higher water/total area ratios than those that were older or post-menopausal. This suggests young and pre-menopausal patients have a higher water content in the non-tumour breast tissue. This is in keeping with previous findings seen with other techniques such as Diffuse Optical Spectroscopy [284] and Magnetic Resonance imaging [285], however, in these studies the optical technique findings were not correlated with the histopathological findings but on radiological imaging. The micro spectroscopy analysis of non-tumour tissue demonstrated that younger patients had areas of stromal tissue (protein rich tissue) which gave a protein peak in the HWN region. However, the HWN spectra obtained from non-tumour tissue was dominated by a lipid peak, rather than a protein peak. The areas of protein tissue are still embedded within surrounding lipid tissue and the sampling area of the probe may include both areas. The Raman cross section of lipid is

greater than protein and only a small amount of lipid is needed within that sampling area to give a lipid peak, which dominates the protein peak in the CH stretch region (similar to the findings with the mixed gelatine/soya bean phantoms in Chapter 3). This results in no observable protein peak and the higher amount of water still measurable in an increased water peak. This would result in a higher water/total area ratio, and perhaps explains the differences observed between non-tumour specimens.

These differences in the physiological 'normal' could account for any diagnostic inaccuracy seen in the method, as at the extremes of variation, the physiological 'normal' spectra overlap with the tumour spectra. In a number of studies investigating the changes in water content to differentiate between tumour and non-tumour breast tissue using optical techniques other than Raman, the range of the physiological 'normal' breast water measurements have overlapped with the range of the 'tumour' water measurements, thus affecting the diagnostic accuracy of using water to differentiate between tissue types [276, 289, 291]. A limitation of these studies was that in the calculation of the 'tumour' water measurements they used information from scans of large areas of breast tissue that contained both tumour and non-tumour tissue, which may have underestimated the degree of difference in water content between tissue types.
If the variation in physiological 'normal' was significant and the cause of diagnostic inaccuracy it may mean that using water content as a main diagnostic marker is less accurate in the young/pre-menopausal in particular. It would therefore be expected that accounting for this variation would improve diagnostic accuracy of the technique. Calculation of the W/TAR difference allowed for interpretation of the tumour Raman spectra to be performed as a measure against the patient's own baseline, provided in a 'calibration' measurement. Therefore, if a patient had a high water/total area ratio baseline in their non-tumour tissue, this would be accounted for as the tumour tissue readings would be measured as any change from that baseline. The results from this analysis did not show an increased diagnostic accuracy compared to other methods of analysis. This suggests that the physiological variation of non-tumour tissue water/total area ratio between patients does not adversely affect the diagnostic accuracy of differentiating between tumour and non-tumour tissue.

4.4.3.3 Lobular Carcinoma

Lobular carcinoma can be a diagnostic challenge; mammography has poor sensitivity in detection of the disease [361], and is poorly circumscribed histologically with a single line infiltration of malignant cells [25]. It may also present a diagnostic challenge in Raman spectroscopy; previous studies using Raman spectroscopy for IMA do not investigate lobular carcinoma as a possible pathological sub-group that may cause diagnostic inaccuracy [189, 262].

The micro spectroscopy analysis of the tumour specimen of patient 14 (which was a lobular carcinoma) demonstrated lipid signals (which are generally associated with non – tumour tissue) even within tumour. It was also noted by the histopathologist on review of the H+E slide, that it was a difficult case to examine and the scattered cancer cells made it a diagnostic challenge.

The histopathological features are different between lobular and ductal carcinoma, and the water/total area ratio was lower for lobular carcinoma, suggesting there are some spectral differences between the two carcinoma types. However, despite this possible diagnostic challenge in the subgroup analysis of all the tumour specimens, although the water/total area ratio was different between ductal (mean ratio - 0.77) and lobular carcinoma (mean ratio. 0.70), it did not reach statistical significance. Comparing these ratios with the mean water/total area ratio for the non-tumour specimens which was 0.2 demonstrates that the variation between carcinoma types was relatively small, compared to the differences between tumour and non-tumour tissue. This suggests that although lobular carcinoma may present a diagnostic challenge for Raman spectroscopy, it is the nature of the pathology, and NP3 is capable of differentiating lobular carcinoma tissue from non-tumour tissue.

4.5 Conclusion

This study has demonstrated the ability of NP3 to differentiate between tumour and non-tumour breast tissue with an excellent diagnostic accuracy of up to 93%. A wide range of data processing techniques using the fingerprint region, HWN region and the water/total area ratio are capable of differentiating tissue types. The water/total area ratio method is simple and could reduce data processing

time whilst providing accuracy equal to more complex spectral analysis, suggesting this may be the most clinically relevant method of data analysis to provide IMA.

The biochemical differences between tumour (protein and water rich) tissue and non-tumour (lipid rich, water poor) have been demonstrated with spectral assignation, and validated with micro spectroscopy Raman measurements, that matched histopathological features with Raman spectral features. The impact of pathological subtypes that have previously been difficult to diagnose, such as DCIS only tissue, the broad range of variation in physiological 'normal' non-tumour breast tissue and lobular carcinoma on the diagnostic ability of the technique has been discussed. Although more work is required, it suggests that the NP3 system is capable of good diagnostic accuracy even in these traditionally more difficult pathological subgroups.

There are some limitations to the study, such as the study was conducted on fresh frozen tissue and the time taken for each spectral measurement was 25 seconds. Further work is required to validate these study findings on fresh breast tissue and a probe and study protocol that could provide IMA within clinically relevant times needs to be developed. However, the findings and analysis of the data from this study suggests how this may be accomplished, and highlights which areas need particular attention.

This study has been successful in providing the proof of principle that NP3, and in particular HWN Raman spectroscopy, can differentiate between tumour and non-tumour tissue with excellent diagnostic accuracy. This gives a basis for further work to translate these findings to providing IMA.

CHAPTER 5:
An optimised Raman system for Intraoperative Margin Analysis

5.1 Introduction

Chapter 4 demonstrated the proof of principle that tumour can be differentiated from non- tumour breast tissue using the NP3 system. However, a number of limitations of the NP3 system were recognised in the acquisition of the data.

Particular problems encountered with the NP3 system were:

- The probe was physically fragile and susceptible to breaking if incorrectly handled or dropped
- Some of the collection fibres did not work, and these were inaccessible/difficult to replace
- Transporting it between the laboratory and the hospital site was difficult (due to fragility)
- A small sampling volume
-

The NP3 system was unlikely to be successfully utilised without further probe manufacturing to become a system capable of providing IMA due to these limitations. To effectively translate the promising findings in chapter 4 towards providing Intraoperative Margin Analysis (IMA), the limitations of the probe need to be addressed. Therefore, using the knowledge gained through the thesis with the NP3 system, and using the same basic set up of a 785 nm laser excitation wavelength and InGaAs camera, a new handheld probe Raman system was developed for future clinical use.

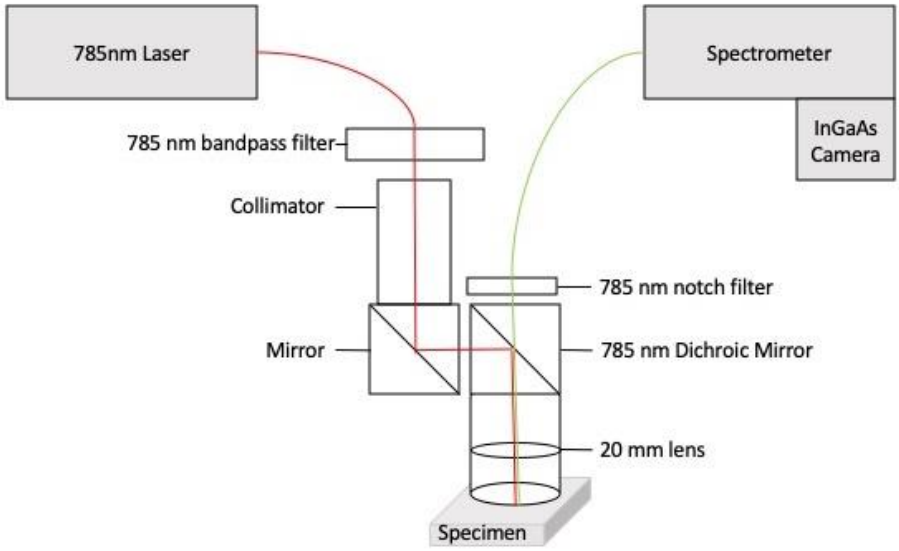
This chapter explores this new handheld Raman system (HH Raman system), ensuring the basic characteristics (i.e. ability to obtain HWN Raman measurements) are similar to the NP3 system, and going on to investigate the limitations and clinically relevant capabilities of the system to assess its suitability to provide IMA in the future.

5.2 Methods

5.2.1 HH Raman system configuration

The HH Raman system is similar in principle to the NP3 system, however the probe delivery system is different (Figure 5-1). Laser excitation was delivered at 785 nm IPS spectrum stabilised laser module (Innovative Photonic Solutions,USA) with a 105µm core fibre through a collimating lens and then a 785nm laser bandpass filter (Thorlabs,NJ,USA), and collimated light passed through a 785 nm dichroic mirror (Thorlabs,NJ,USA). A plano-convex 20mm lens with anti-reflective coating (Thorlabs,NJ,USA) and a 20mm spacer ensure light is focused at the tip of the probe, providing a maximum power of 250mW to the sample with a 0.19mm spot size. The light was then collected through an SMA 7X 105 µm collection bundle with a collection spot size of 0.64 mm through a 785nm notch filter to the entry port on the spectrometer - Kaiser Holospec Imaging Spectrograph (Kaiser optical systems inc, Ann Arbour,USA) with a broad HVG 800 grating which is coupled to a camera InGaAs camera (iDus InGaAs 1.7µm, Andor, Belfast, UK) cooled to -70°C.

A



B

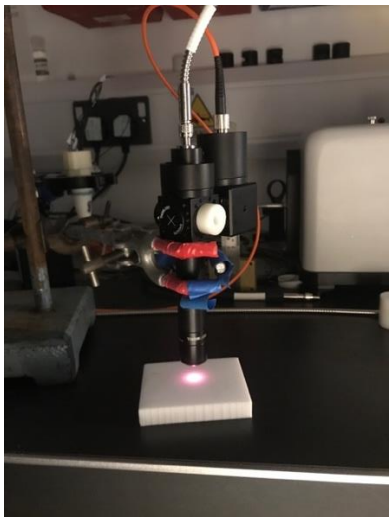


Figure 5-1 HH Raman system configuration

A. Schematic diagram showing the system components, configuration and light path. Red line is excitation light (785 nm), Green line is Raman scattered light. B. Picture of handheld component taking a measurement from block of PTFE. The spot looks much larger than it is due to the significant scattering in the PTFE block.

5.2.2 Breast Phantom construction

To assess the HH Raman system, and to compare the results with all previous work in the thesis, the same phantoms were used for testing purposes. Gelatine phantoms were made to produce a range of water concentrations 85-95% using the same methods described in full in Chapter 2.

Pork tissue was used to mimic biological tissue, analogous to fat rich, normal breast tissue (pork fat, and pork lard in later lumpectomy phantom construction) and protein and water rich tumour tissue (pork meat). Pork tissue and lard were purchased from a local supermarket (Tesco, UK).

The composition of each phantom was specific to each experiment, and is described in the corresponding section below.

5.2.3 Measurement protocol

Spectra were recorded using SOLIS software (Andor, UK) and processed in Matlab (Mathworks, USA). Daily calibrations were performed with Neon Argon lamp, Ethanol, water, PTFE, paracetamol and aspirin. The Raman peaks of paracetamol in the HWN region were used to calibrate the Raman shift using the inbuilt calibration tool in the SOLIS software. A dark signal measurement (with laser and background lights off) that corresponded to the signal acquisition time was taken daily for subsequently detailed spectral pre-processing.

Specimens were wrapped in clingfilm to reduce probe contamination, and because in clinical studies the probe will need a disposable protective sheath of clingfilm for each measurement. Measurements taken of pork tissue with and without clingfilm on the surface revealed no discernible difference to the resulting Raman spectra – data not shown. As a main aim of these experiments was to reduce signal acquisition time - signal acquisition for these experiments was a single acquisition of 5 seconds per measurement (apart from in the experiments evaluating decreasing signal acquisition time).

Measurements were taken with the HH Raman system, with the probe in light contact with the specimen (apart from in the experiments evaluating increasing this distance). 5 measurements were taken, at random, over the surface of each phantom.

5.2.4 Data processing and analysis

Spectra were pre-processed initially by dark noise subtraction. A smoothed dark noise baseline was obtained by a Savitzky-Golay filter with a 7^{th} order polynomial of the dark noise reading of the corresponding spectral acquisition time (5 seconds unless otherwise stated in the individual sections) and this was subtracted from each spectrum. Baseline subtraction was undertaken with a 2^{nd} order polynomial fit. Spectra presented are the mean spectra of the 5 repeats (n=5) +/- 1 Standard Deviation (SD), and have been normalised to the CH stretch region between 2850-2950 cm^{-1}. Each spectrum was analysed using the

water/total area ratio method, (area under the curve of 3035-3680 cm^{-1} region divided by area under the curve of 2850-3680 cm^{-1} region) as previously described in Chapter 3 and 4. The mean water/total area ratio +/- 1 SD is used to describe each specimen. Scatter graphs were constructed for the gelatine phantoms using this data plotted versus the known water fraction of the gelatine phantom, and a linear fit calculated based on all points with quality of fit described by the Root Mean Square Error (RMSE).

Statistical analysis to compare water/total area ratio was performed with one - way ANOVA, statistical significance $P<0.05$.

5.3 Experimental protocol and results

5.3.1 Assessment of probe detecting changes in water content

The initial test of the HH system was to ensure it could obtain Raman spectra with a good signal to noise ratio using the phantoms that had previously been used to mimic breast tissue (pork meat and fat), as these were to be used in subsequent experiments. Figure 5-2 shows that high quality, high signal to noise, Raman spectra were obtained with just 5 second acquisition times.

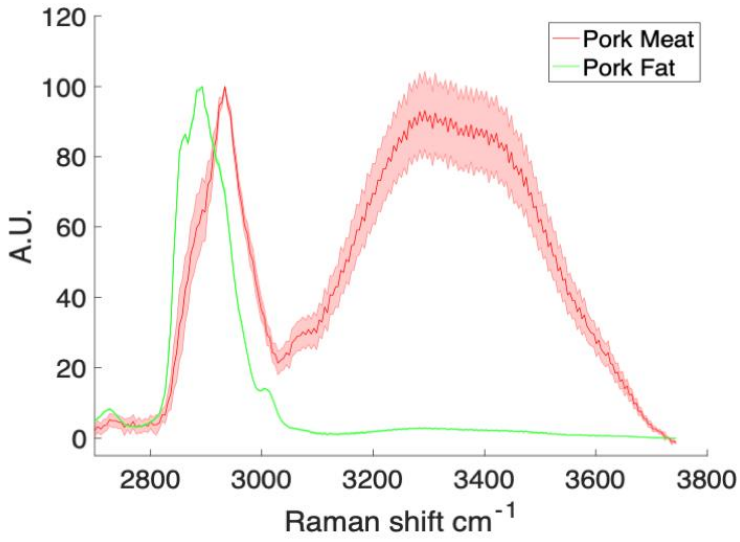


Figure 5-2 Showing Raman spectrum of pork meat and pork fat obtained with HH Raman system. Spectra are mean spectra with shading in the same colour either side +/- 1 S.D. (n=5). Acquisition time 5 seconds.

The ability of the HH Raman system to detect changes in water content were then tested. The results from the fresh frozen breast tissue suggest that normal, fatty,

breast tissue has a very low water content and that changes in the water content cannot be measured using HWN Raman – therefore changes in water content in a lipid environment were not tested as it is of limited clinical relevance. However, the protein rich tumour areas had a range of water contents that could be measured using the water/total area ratio calculation. Therefore, gelatine phantoms at a range of water concentrations (95-85% water) were measured to test if the HH Raman system could measure changes in water concentration using the water/total area ratio method of spectral analysis. Figure 5-3 shows that there was a linear relationship between known water fraction and the water/total area ratio, with a gradient of 0.533 and RMSE of 0.002. This demonstrates that the HH Raman system can obtain HWN spectra of sufficient quality to detect small changes of 2.5% (in a protein rich environment) in water concentration. These results are similar to the results from the NP3 system (where similar plots of water fraction vs water/total area ratio had a gradient of 0.57), suggesting that the HH Raman system has comparable performance in measurement of the HWN region.

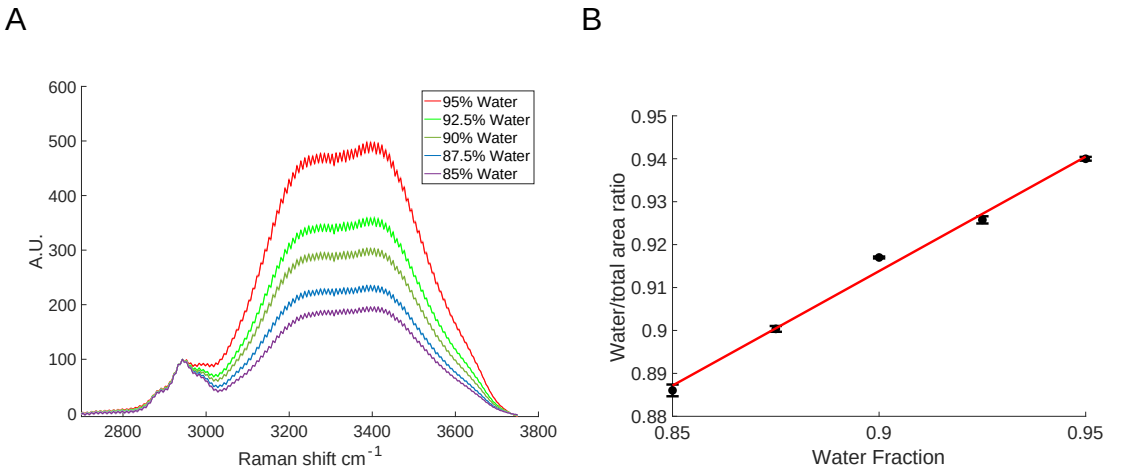


Figure 5-3 showing Raman spectra in the HWN region of different water concentrations of gelatine phantoms and the relationship with the water/total area ratio

A. Graph showing Raman spectra of gelatine phantoms at 5 different water concentrations, normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each concentration after pre-processing, shading in the same colour either side is +/- 1 SD. B. Scatter graph of known water fraction versus water/total area ratio. Points plotted are mean water/total area ratio (n=5), error bars +/- 1 SD. Red line is line of best fit (Gradient 0.53, RMSE = 0.002)

5.3.2 The effect of changing measurement parameters on signal acquisition

5.3.2.1 Decreasing laser power

The standard power of the 785 nm laser for these experiments was 250 mW. Higher laser powers could be obtained and attempted with this system (up to 350 mW), but this led to sample degradation through burning. With the power at 250 mW no burning of any samples was observed, this was also the power that was used for the NP3 system measurements of fresh frozen breast tissue (where the power density is much higher due to the laser being focused on a smaller area [362]) where there was no burning of small, dry, biopsy samples. The method of IMA that is envisaged with the HH probe is for *ex vivo* analysis of excised breast tissue samples. Therefore, although burning of samples is to be avoided as it may confound subsequent histopathological analysis, it would not cause immediate harm to a patient. It may be that future uses for the HH Raman system could be *in vivo* assessment of the tumour bed, similar to work done in neurosurgery [223]. It is possible that the laser power would have to be reduced below 250 mW to avoid deleterious tissue temperature increases and burning for future *in vivo* studies. Most studies using Raman spectroscopy *in vivo* on skin [363] and within the abdominal cavity [248] use a delivered power of 100 -150 mW, and *in vivo* analysis of the tumour bed it was as low as 37-64 mW [223] (though no studies actually measured specimen temperatures/degradation).

As the intensity of Raman scattering is related to the laser power, reducing laser power could negatively affect signal acquisition [362]. To investigate the potential for the system to be used *in vivo* on biological tissue at reduced power, a series of measurements were performed on pork meat at different laser powers. The Raman cross section of protein and water is much lower than lipid [364], and so any anticipated reduction in signal to noise ratio would affect the ability to detect protein and water more than the ability to detect lipid. Therefore, pork meat was measured as the detection of protein and water rich areas (mimicking a tumour environment) is the most clinically relevant.

Figure 5-4 shows the results of measuring pork meat at different laser powers. At all laser powers there is good signal to noise and Raman signal could be recovered. Analysis of the water/total area ratio between the laser powers demonstrates that the mean ratio was different between laser powers (250 mW

– 0.794 (SD 0.006), 150 mW – 0.787 (SD 0.006), 80 mW – 0.776 (SD 0.012)) and this reached statistical significance (P= 0.016) and multi-group comparison of the ratios was significant between 250 mW and 80 mW (P= 0.013), but not between any other group comparisons (250 mW vs 150 mW; 150 mW vs 80 mW; P>0.05).

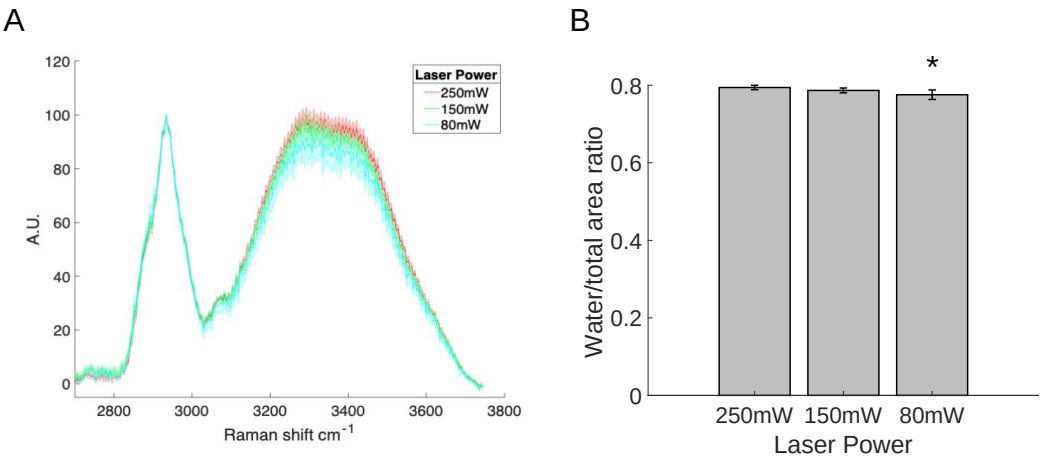


Figure 5-4 Results of measurement of pork meat with HH Raman system with different laser excitation powers

A Graph showing Raman spectra of pork meat at 3 different laser powers, normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each concentration after pre-processing, shading in the same colour either side is +/- 1 SD. B. Bar graph comparing the mean water/total area ratio between the signal acquisition times, error bars +/- 1 SD. One-way ANOVA to compare between samples, comparison between 250mW and 80mW revealed statistical significance, denoted by * P=0.013

These results demonstrate that the HH Raman system can be used with a decreased laser power, and Raman signal is obtained at the low power of 80 mW. Although the water/total area ratio is statistically different between the laser powers, in real terms it only varies maximally by a ratio value of 0.04 between the highest and lowest power, and this would be unlikely to adversely affect the diagnostic ability in differentiating tumour and non- tumour tissue (in the fresh frozen samples the differences in water/total area ratio between tissue types were in the order of 0.5). This may be useful if future work requires the probe to be used at lower laser powers for *in vivo* use.

5.3.2.2 Measurement acquisition time

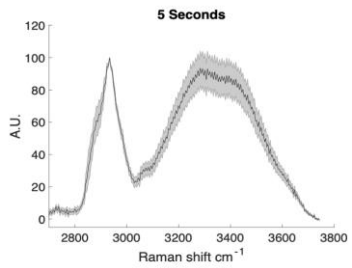
A limitation in previous experiments in this thesis has been the acquisition time to obtain measurements. In the fresh frozen breast specimens, a total acquisition time of 25 seconds was used to ensure adequate signals were obtained for the

proof of principle phase. Future work with the HH Raman system will be in the clinical environment and reducing the signal acquisition time would allow a reduction in time for specimen analysis.

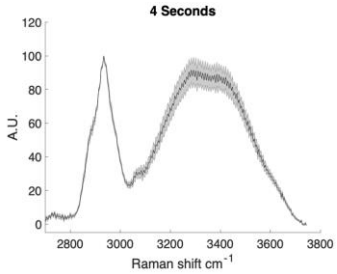
A limitation of InGaAs cameras is the increased dark noise, requiring additional pre-processing steps, and readout detector noise, requiring longer acquisition times to overcome this, compared to CCD cameras [340, 365]. Santos *et al.* used an InGaAs camera in biological tissue (skin) and used acquisition times of 10 seconds, and when reduced acquisition times were attempted the results were poor (though they were measuring pigmented skin lesions) [365]. This is in comparison to a recent study by Liao *et al.* a CCD camera and excitation wavelength of 671 nm with a 2-axis scanning galvo- mirror system was used to obtain signals with a signal acquisition time of 10 ms, however, this did not provide signals in the entire HWN region (omitting the water peak), there was significant background fluorescence and the HWN signals were not diagnostic signals [299]. This work by other groups suggests that advancing the HH Raman system and the use of InGaAs cameras requires the ability to reduce signal acquisition time to compete with the timescales achieved with CCD cameras for signal acquisition.

In the previous sections in this chapter, an acquisition time of 5 seconds has been used, and the results have so far demonstrated that this provides excellent Raman signals with good signal to noise, and that this reduced acquisition could be used in the future. Reducing the signal acquisition times even further could have great clinical benefits in progressing the HH Raman system closer towards providing IMA in clinically relevant timeframes. Therefore, a series of measurements were obtained on pork meat over a range of signal acquisition times (from 5 seconds to 0.1 seconds), and the resultant spectra water/total area ratio compared to determine the effect of signal acquisition times on obtaining spectra in the HWN region (Figure 5-5).

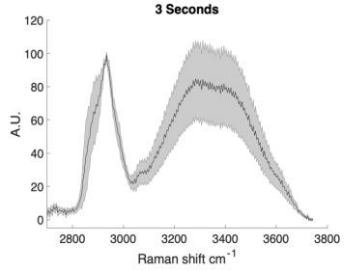
A


5 Seconds
A.U.
Raman shift cm-1

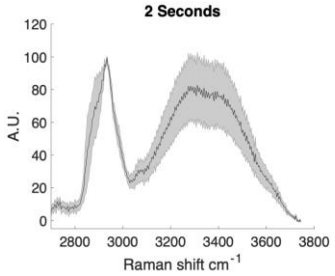
B


4 Seconds
A.U.
Raman shift cm-1

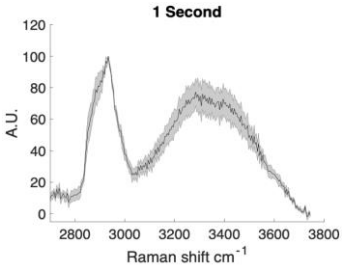
C


3 Seconds
A.U.
Raman shift cm-1

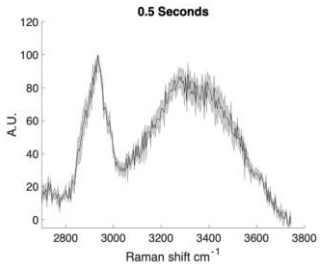
D


2 Seconds
A.U.
Raman shift cm-1

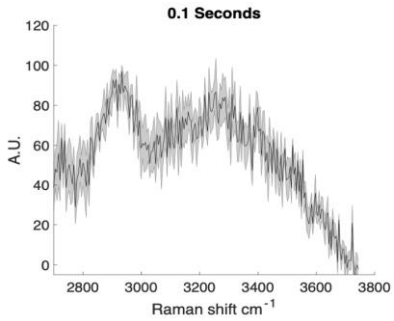
E


1 Second
A.U.
Raman shift cm-1

F


0.5 Seconds
A.U.
Raman shift cm-1

G


0.1 Seconds
A.U.
Raman shift cm-1

H

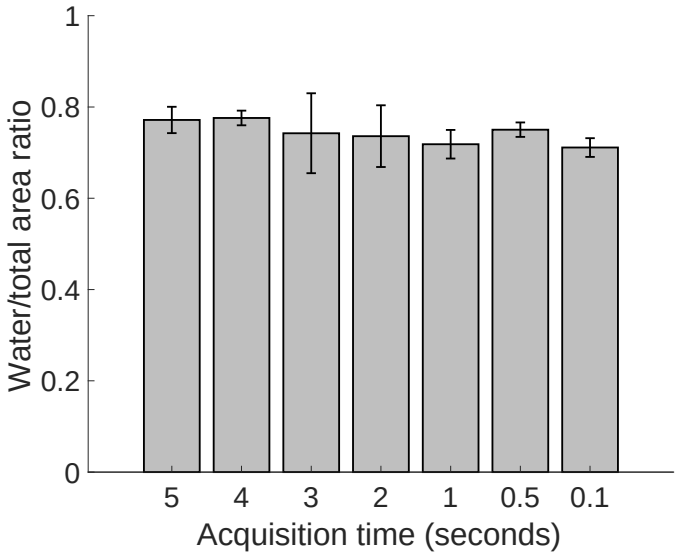

Water/total area ratio
Acquisition time (seconds)
5
4
3
2
1
0.5
0.1

Figure 5-5 Raman spectra of pork meat obtained with HH Raman system with a range of signal acquisition times (5 second – 0.1 second)

A – G –Raman spectra of pork meat at different signal acquisition times (titled), normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each acquisition time after pre-processing, shading in the same colour either side is +/- 1 SD. H. Bar graph comparing the mean water/total area ratio between the signal acquisition times, error bars +/- 1 SD. One-way ANOVA to compare between samples, found no significant difference between water/total area ratio P>0.05

Figure 5-5 shows that discernible Raman signal was recovered at all signal acquisition times, even to the very short time of 0.1 second. The spectra for acquisition times of 0.5 seconds and greater were remarkably similar in appearance. However, spectra obtained at 0.1 seconds are visually poor quality, with a low signal to noise, and the background noise impacting significantly on the spectra particularly in the region of the OH stretch (3035 – 3680 cm^{-1}). Analysis of the water/total area ratio for the spectra showed the mean water/total area ratio did decrease with signal acquisition time (from 5 seconds – 0.77 to 0.1 second - 0.71) but this did not reach statistical significance on multiple comparison between all groups (P >0.05).

These signal acquisition times are considerably shorter than other Raman systems utilising an InGaAs camera, while still maintaining Raman spectra with reasonable signal to noise [345, 365]. One reason could be the pre-processing that we perform with dark noise subtraction, which is not performed in the other studies. The quality of Raman spectra obtained at 0.1 second (100 ms) is comparable, if not better (considering they are with less scattering protein rich tissue), than those obtained in fatty tissue with a CCD at 10 ms [299]. Some biochemical information (the water peak) may be lost at these lower acquisition times, however, depending on the spectral features required for analysis, acquisition times could be reduced further.

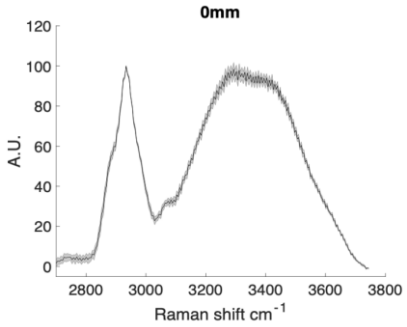
These results demonstrate that it is possible to reduce signal acquisition times to as low as 0.5 seconds with the HH Raman system and our pre-processing methods and obtain sufficient quality Raman spectra for full biochemical analysis and interpretation.

5.3.2.3 Reduced probe contact

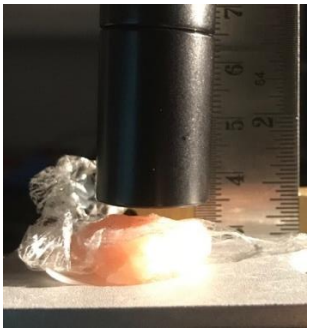
An element of signal acquisition that is particularly important with a handheld probe is the variability in contact between the probe and the surface being measured. These variations may occur either due to human factors (i.e. not consistently putting the probe directly in contact with the specimen) or specimen factors (i.e. unequal surface of specimen). This occurred in the use of the ClearEdge probe which caused an error in the device and the margin could not be assessed in a number of specimens [266]. This is a potential major limitation with a handheld probe.

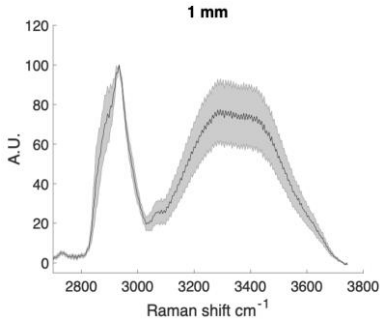
The HH Raman system is designed for the probe objective to be in light contact with the surface of the specimen throughout the acquisition time, as the 20 mm lens focuses the excitation light at that point. However, it is important to assess the effect that variation would have if the specimen was measured not in the 'ideal conditions' of having the probe in light contact with the specimen, as this is likely to happen with some measurements in a clinical setting. A series of measurements were taken with the probe at a range of distances from the surface of the specimen (pork meat) to investigate the effect that reduced probe contact with the specimen has on obtaining Raman spectra (Figure 5-6).

A

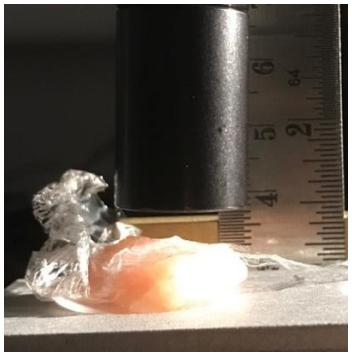

0mm
A.U.
Raman shift cm-1

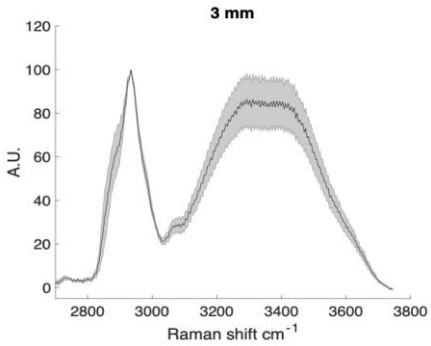
B



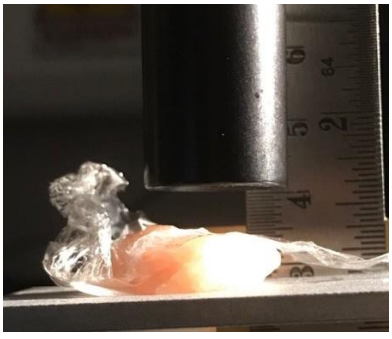

1 mm
A.U.
Raman shift cm-1

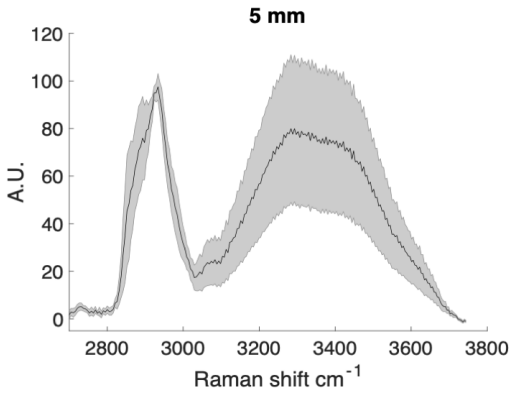
C



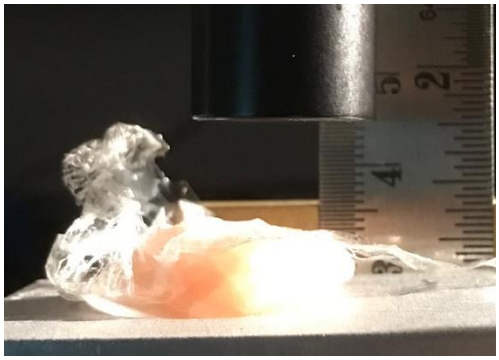

3 mm
A.U.
Raman shift cm-1

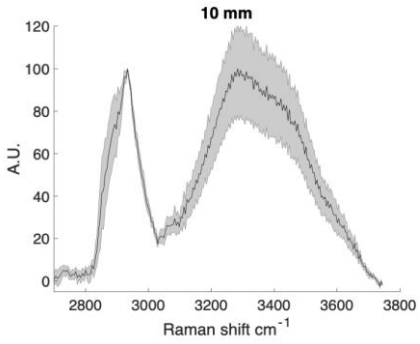
D




5 mm
A.U.
Raman shift cm-1

E





F

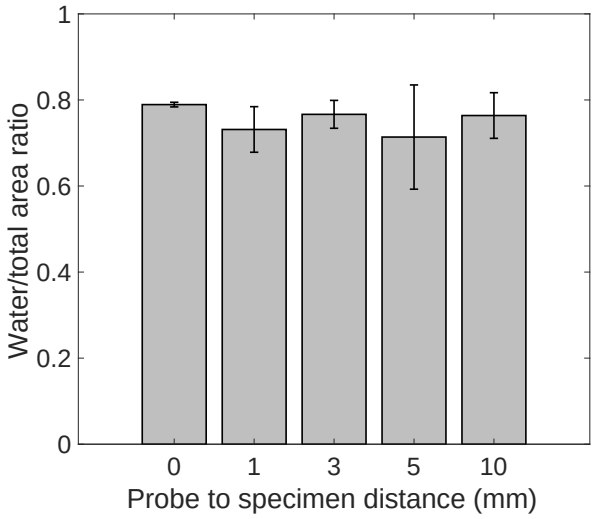


Figure 5-6 Raman spectra of pork meat obtained with HH Raman system with a range of probe to specimen distances (0 mm – 10 mm)

A – E –In the left column is a photo demonstrating the distance from probe to specimen (pork meat in clingfilm), in the right column is Raman spectra of pork meat at different probe to specimen distances (titled), normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each acquisition time after pre-processing, shading in the same colour either side is +/- 1 SD. H. Bar graph comparing the mean water/total area ratio between the probe to specimen distances, error bars +/- 1 SD. One-way ANOVA to compare between samples, found no significant difference between water/total area ratio P>0.05

Figure 5-6 shows the spectra obtained of pork meat at a range of probe to specimen distances. It can be seen that the spectra obtained when the probe is in contact with the specimen (0 mm probe to specimen distance) has a smaller standard deviation by a factor of 10 compared to other measurements (mean water/total area ratio (SD)- 0 mm 0.789 (0.005); 1 mm 0.731 (0.053); 3 mm 0.767 (0.031); 5 mm 0.714 (0.121); 10 mm 0.764 (0.053)). This suggests the least variability in measurements when the probe is in contact with the specimen.

However, when the water/total area ratios were compared between all groups they were not significantly different (P>0.05).

There are a number of reasons why introducing a gap between the probe and specimen would introduce greater spectral variability, such as stray light [263] and reduced efficiency of collection of Raman scattering [160]. These results suggest the optimal method of measurement is with the probe in light contact with the specimen, but that adequate Raman spectra can be obtained when the probe is up to 10 mm away from the surface of the specimen. Therefore, if there was deviation from measurement protocol, and the probe takes measurements not in contact with the specimen, it should not critically affect results.

5.3.2.4 Sub-surface Sampling depth

A characteristic of the probe that has particular clinical relevance is the sampling depth. The sub-surface sampling depth of a Raman system is defined by the lateral distance from the illumination source to the collection system [160]. In traditional Raman microspectroscopy, where the illumination and collection light are in the same place, the sampling depth is in the order of a few hundred microns and is predominantly a surface technique [366]. If there is a spatial offset the photons that have travelled some depth into the tissue are collected, thus likely to collect scattered photons from beyond the surface of the sample [160]. There is a clinical significance to this, as if there is disease within 2 mm of the resected edge re-excision should be considered, and if disease is at the resected edge, re-excision should be undertaken [56]. In the HH Raman system the illumination source and collection bundle are not separated and there is negligible spatial offset. Therefore, there should be minimal contribution from spatial scattered photons beyond the sample surface. Understanding the sampling depth of the probe is important to interpret the Raman signals, and what the clinical significance may be, when it is applied to IMA.

Most studies examining sampling depth measure a highly scattering substance at increasing depths [367] to ascertain if the Raman signal from the buried medium can be detected. However, for a clinical probe this is of limited value and significance, as the ability to measure a buried medium is dependent on the Raman cross section of the buried medium i.e. if it is highly scattering the Raman signal will be detected at much greater depths than if it is poorly scattering [160].

Therefore, as the biological tissue will have a different Raman cross section to the experimental buried medium, the sampling depth will not necessarily be valid in the clinical application [368].

To further understand the sampling depth of the probe with clinical relevance, substances that have analogous scattering properties to human breast tissue were used. Pork fat was used to mimic normal, fatty breast tissue as it is highly scattering [162, 314]. For the first experiment measuring fatty tissue underneath a protein rich tissue, a 90% water gelatine phantom was used to mimic protein rich tumour tissue, rather than pork meat. This is because gelatine phantoms have lower scattering properties than pork meat, and it was noted in chapter 4 that tumour tissue had a low degree of scattering (with low intensity signals), and have been used to mimic tumour tissue previously [304]. Gelatine phantoms will have lower scattering than breast tumour tissue, and therefore these models represent a 'worst case scenario' of a low scattering, thin layer of tumour overlying a thick layer of highly scattering normal fatty tissue.

The first experiment measured pork fat with an increasing depth of gelatine phantom on top. The 90% water gelatine phantom was sliced into 1 mm deep strips and placed on top of the pork fat (range from no gelatine– 7 mm of gelatine). This is analogous to measuring a tumour overlying normal fatty tissue.
Figure 5-7 shows that with 1 and 3 mm of gelatine phantom overlying the pork fat, the lipid signal is predominant, though the water content of the gelatine phantom is apparent. With 5 and 7 mm of gelatine phantom the protein and water of the gelatine phantom predominate, though there is still some residual lipid signal contribution to the spectra. Analysis of the water/total area ratio shows that the mean ratio increases with the depth of gelatine over the fat, highlighting the increasing contribution from the water peak to the spectra and these were statistically significant (No Gelatine water/total area ratio – 0.12 (SD 0.015) versus 1 mm Gelatine – 0.41 (SD 0.16); 3 mm Gelatine 0.46 (SD 0.035); 5 mm Gelatine 0.72 (SD 0.054) and 7 mm Gelatine 0.79 (SD 0.031); all comparisons $P=<0.001$).

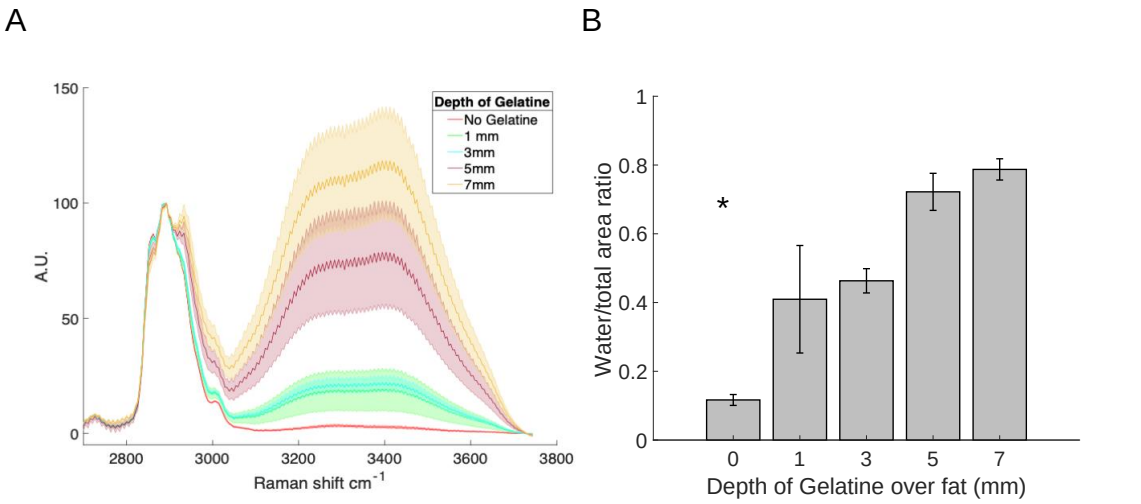


Figure 5-7 Raman spectra of pork meat obtained with HH Raman system with a range of depths of 90% water gelatine phantom overlying the fat (0 mm – 7 mm)

A Graph showing Raman spectra of pork meat with different depths of overlying gelatine phantom, normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each concentration after pre-processing, shading in the same colour either side is +/- 1 SD. B. Bar graph comparing the mean water/total area ratio between the signal acquisition times, error bars +/- 1 SD. One-way ANOVA to compare between samples, comparison between 0 mm (no gelatine) and all other depths of gelatine phantom revealed statistical significance, denoted by * P=<0.001

The second experiment aimed to assess the sampling depth with lipid overlying a protein rich sample. Due to the results from experiment one suggesting that the gelatine phantom has a low Raman cross section, it seemed unlikely that gelatine phantom under a layer of lipid would be detected. Therefore, pork meat (which had been observed to have a more intense Raman signal) was used as the protein rich sample. To get as thin a layer of lipid over the top of the meat, pork lard was used as fat could not be cut thin enough. This model replicates an area of tumour, with a very thin layer of normal fatty tissue overlying it – this could represent a positive margin that requires consideration for re-excision.

Figure 5-8 demonstrates that even with as thin a layer of lard as possible, no meat signal could be recovered, as the lard is so highly scattering the lipid signal dominates.

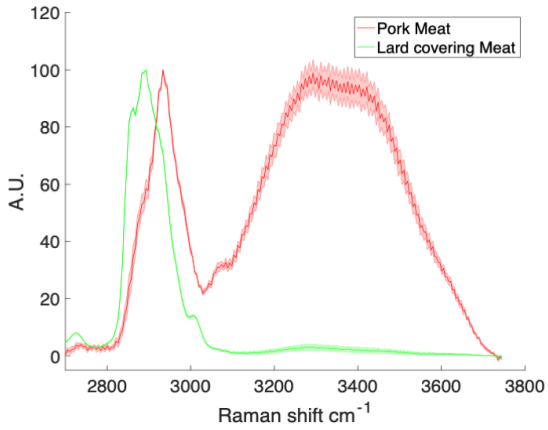


Figure 5-8 Raman spectra of pork meat and lard covering pork meat obtained with HH Raman system
Graph showing Raman spectra normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each concentration after pre-processing, shading in the same colour either side is +/- 1 SD.

These results demonstrate that when a protein rich layer with low scattering and a low Raman cross section (gelatine phantom) is overlying a highly scattering agent (pork fat), significant contribution from the highly scattering agent persists at substantial depths of up to 5 mm using the HH Raman system. However, the water/total area ratio analysis suggests that despite this, the Raman spectra are significantly different between just fat, and fat with gelatine overlying it. In the second experiment protein rich tissue (mimicking tumour) could not be detected beneath a thin (<1 mm) layer of lipid rich tissue (mimicking normal fatty tissue).

There are some limitations to these experiments. The thickness of the gelatine slices in experiment 1, and the thickness of the lard in experiment 2 could not be precisely established. It may also be that the scattering properties and Raman cross section of the phantoms may not exactly mimic breast tissue, which would alter the subsurface sampling depth. The first experiment of gelatine overlying pork fat represents a ‘worst case scenario’ of small amounts of low scattering breast tumour overlying normal fatty tissue. As breast tumour tissue is more scattering than gelatine phantoms, this issue may not be so pronounced with clinical tissue measurements. Despite these limitations, this work highlights the difficulty in detecting protein rich tissue (tumour) below even a thin layer of lipid rich tissue (normal breast tissue). Further work is required when there is access to human breast tissue to ascertain the depths below normal fatty tissue at which tumour tissue can be detected. If this is found to be necessary in clinical

specimens, the probe could be modified to introduce a spatial offset to maximise signal from below the surface.

5.3.3 Assessment of probe in 'tumour' detection

Assessing the sampling area of the probe is an essential step in understanding the capability of the probe and the clinical significance of the results generated from the HH Raman system. The important clinical question is – what is the smallest area of tumour that can be detected by the probe?

Raman based techniques have been reported to detect residual tumours on the surface of specimens as small as 1 mm^2, though the diagnostic threshold sensitive enough to detect these smaller tumours had a specificity of 83% indicating a significant number of false positives using this diagnostic threshold[189]. The MarginProbe (currently the most advanced and 'at market' tool for IMA) reported a sensitivity of only 56% for detecting tumours of 0.7mm [369]. These literature reports suggest that detection of tumours around 1 mm^2 is necessary for the HH Raman system to be clinically useful, and competitive with other devices and technologies.

Similar to the previous work on investigating sampling depth, experimental methods with other materials could ascertain, perhaps with greater theoretical precision, the sampling area of the probe. However, as the diameter of the sampling volume is affected by the optical properties (such as refraction, scattering, absorption) of the measured specimen, using biological tissue gives a closer understanding of what may be seen with human breast tissue in the clinical setting [362].

For these experiments a series of lumpectomy phantoms were created from pork lard (to mimic normal, fatty breast tissue) and pork meat (to mimic protein and water rich tumour breast tissue). Pork meat was dissected into pieces ranging from 0.5 mm^2 to 5 mm^2 and placed into squares of pork lard (Figure 5-9).

A

B

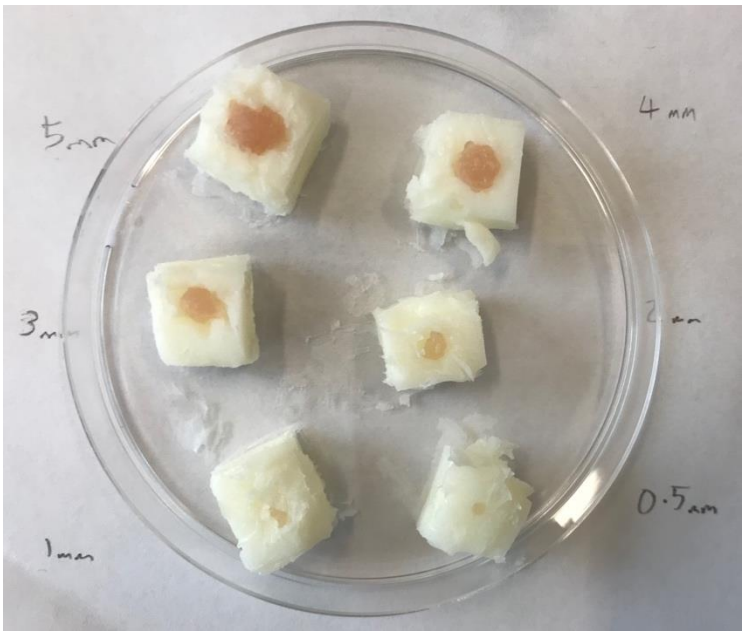


Figure 5-9 Figure showing the lumpectomy phantoms

A. pork meat (tumour mimic) cut into a range of sizes (0.5-5mm^2). B. Lumpectomy phantoms with the pork meat (tumour mimic) of a range of sizes (written next to the corresponding phantom) in a square of lard (normal breast tissue mimic)

5.3.3.1 Point scans

To assess the sampling area, measurements were taken directly over the area of pork meat of each lumpectomy phantom. If the sampling area is greater than the piece of meat, it would also sample the surrounding lard. As lipid is much more scattering than protein, the lipid signal may dominate the resultant spectra and the protein and water signal may not be detected. This would result in that 'tumour' size not being detectable. Therefore, the size of the pork meat in lard ranged from as small as possible that was able to be dissected and measured accurately (0.5 mm^2) and was predicted to be smaller than the sampling area, to

large pieces (5 mm^2) that, if remained undetectable with the probe, would render the probe clinically inappropriate.

The results displayed in Figure 5-10 A demonstrate the change in the Raman spectra as the pork meat ('tumour') increases in size. As there is an increase in size of the pork meat, the protein signal and the water peak become more dominant in the spectrum. The water peak is small in the 0.5 mm lumpectomy specimen which could be due either to a small amount of water in the pork meat (with rapid dehydration from the surface due to relatively large surface area), poor targeting of the probe on the lesion, or, more likely, because of dominant lipid signals giving a relative decrease in the water peak. As there is some lipid contribution to this spectrum, this suggests that the sampling area of the HH Raman system is at least 0.5 mm^2. This would likely adversely affect the diagnostic ability of the system in tumours of this size.

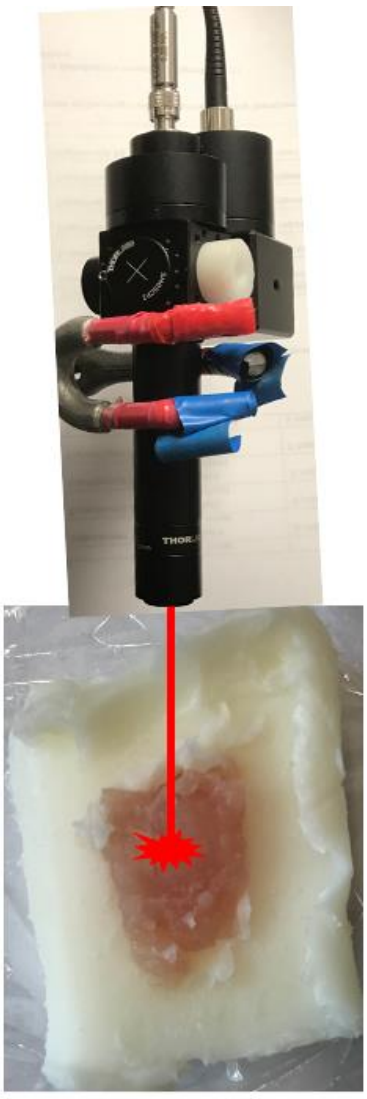
The water/total area analysis in Figure 5-10 B shows that there is a difference in the mean water/total area ratio between pure lard (0.013 (SD 0.011)) compared to the lumpectomy phantoms (0.5 mm phantom 0.31 (0.021); 1 mm phantom 0.59 (0.0055); 2 mm phantom 0.51 (0.068); 3 mm phantom 0.49 (0.056); 4 mm phantom 0.60 (0.036); 5 mm phantom 0.67 (0.030)), and this reached statistical significance ($P<0.001$; one-way ANOVA). All lumpectomy phantoms had significantly different water/total area ratios to each other (1 mm phantom versus 2 mm phantom ($P=0.03$) and 3 mm phantom ($P=0.0026$), all other comparisons $P<0.001$), apart from 1 mm phantom and 4 mm phantom ($P=0.99$) and 2 mm phantom and 3 mm phantom ($P=0.99$).

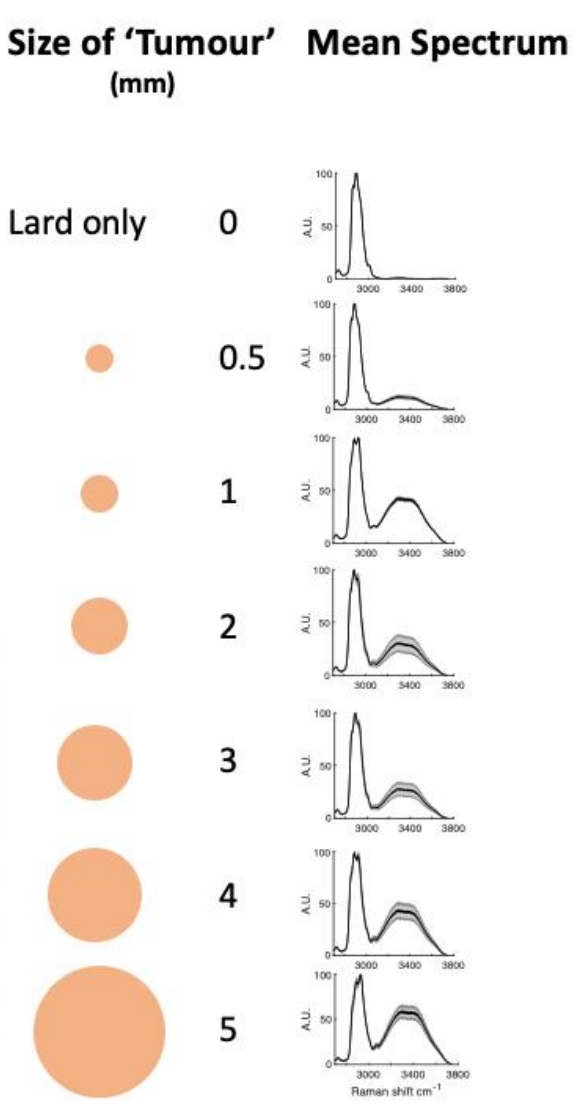
The limitations to these experiments are that the pork meat may have a different Raman cross section to real human breast tumour tissue, which may alter the detection rate. The lumpectomy phantoms with >1 mm^2 pieces of pork meat had different water/total area ratios on comparison, which suggests there may have been other influences on the Raman signal than just the pork meat (i.e. the sampling area may have been greater than 1 mm^2). However, the water/total area ratios for these phantoms were of a similar magnitude (0.49 - 0.67), compared to that of lard (0.013). Therefore, it is more likely that this variation represents natural variation in the pork meat used, as a different piece of pork meat was

used for each lumpectomy phantom, which will have had slightly different biochemical/water properties, and a different rate of dehydration, all of which will affect the water/total area ratio.

These results demonstrate that all lumpectomy phantom ‘tumours’ (ranging from 0.5 – 5 mm^2) had significantly different water/total area ratios compared to lard (mimicking normal breast tissue). This suggests that tumours as small as 0.5 mm^2, and certainly tumours of 1 mm^2, could be detected with the HH Raman system.

A




Size of 'Tumour' (mm)
Mean Spectrum
Lard only
0
0.5
1
2
3
4
5
A.U.
Raman shift cm⁻¹

B

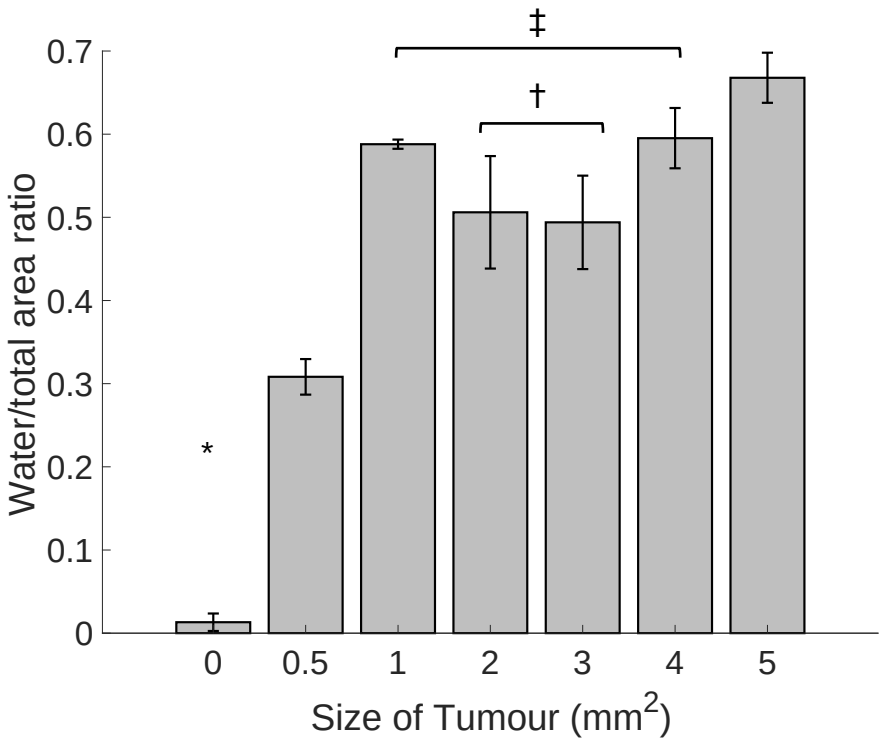

Water/total area ratio
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
*
†
‡
0
0.5
1
2
3
4
5
Size of Tumour (mm²)

Figure 5-10 Raman spectra of lumpectomy phantoms obtained with HH Raman system with a range of pork meat (to mimic 'tumour') sizes (0 mm^2 – 5 mm^2)

A. On the left column is a graphic illustrating the measurements were taken from the centre of the 'tumour' in the lumpectomy phantom. The lumpectomy phantoms had a range of pork meat areas to mimic different tumour sizes. On the right column is the resultant Raman spectra from the centre of the 'tumour' for the corresponding lumpectomy phantom, normalised to CH stretch region between 2850 – 2950 cm^{-1}. Plotted lines are mean (n=5) for each acquisition time after pre-processing, shading in the same colour either side is +/- 1 SD. B. Bar graph comparing the mean water/total area ratio between the lumpectomy phantoms, error bars +/- 1 SD. One-way ANOVA to compare between samples, found all water/tumour area ratios were significantly different between the phantom with no pork meat (0mm) and all other lumpectomy phantoms (denoted by * $P<0.001$) and all lumpectomy phantoms were significantly different from each other ($P<0.05$), apart from between lumpectomy phantoms with 'tumours' of 2mm and 3mm (denoted by †= $P>0.05$), and 1mm and 4mm (denoted by ‡= $P>0.05$).

5.3.3.2 Line scanning

The final component of assessment of the HH Raman system was to conceptualise how the system may be used in the future to provide IMA for Breast Conserving Surgery. One method would be using the probe as a hand-held device and taking measurements of an area of concern – similar to the method described above. However, this would rely on the surgeon identifying an area of concern for analysis, and would not be capable of scanning an entire resected edge / specimen in a methodical manner. A way of measuring an entire resected edge systematically would be to have the probe stationary, and the specimen on a mechanical stage that would automatically move the specimen in the appropriate step sizes for each measurement to ensure the entire area of the resected edge was measured. This could then give a real time read out of the specimen as it is scanned, or a binary outcome at the end of scanning of whether the resected edge is clear or not of tumour.

To test the feasibility of this potential use, the HH Raman system was kept stationary, and a lumpectomy phantom was moved manually with a roughly 1 mm step size in a line through the tumour. The lumpectomy phantom with a 1 mm^2 piece of pork meat in the centre was measured, to mimic a small 1 mm^2 tumour at the resected edge of a lumpectomy specimen. The signal acquisition time was 1 second, to further reduce the time taken for specimen analysis. The acquired spectra were pre-processed and the water/total area ratio calculated. In clinical practise for IMA the surgeon would want a binary, and easily interpretable, output

of the Raman spectra findings to determine whether the margin was positive or not and aid their clinical decision to resect further tissue or not. Therefore, based on the results and methods in chapter 4 (using binomial logistic regression to obtain the optimal water/total area ratio cut off to detect tumour) a water/total area ratio of 0.5 was used to classify the resulting spectrum as tumour (water/total area ratio ≥ 0.5; red) or normal (water/total area ratio <0.5; green) (Figure 5-11).

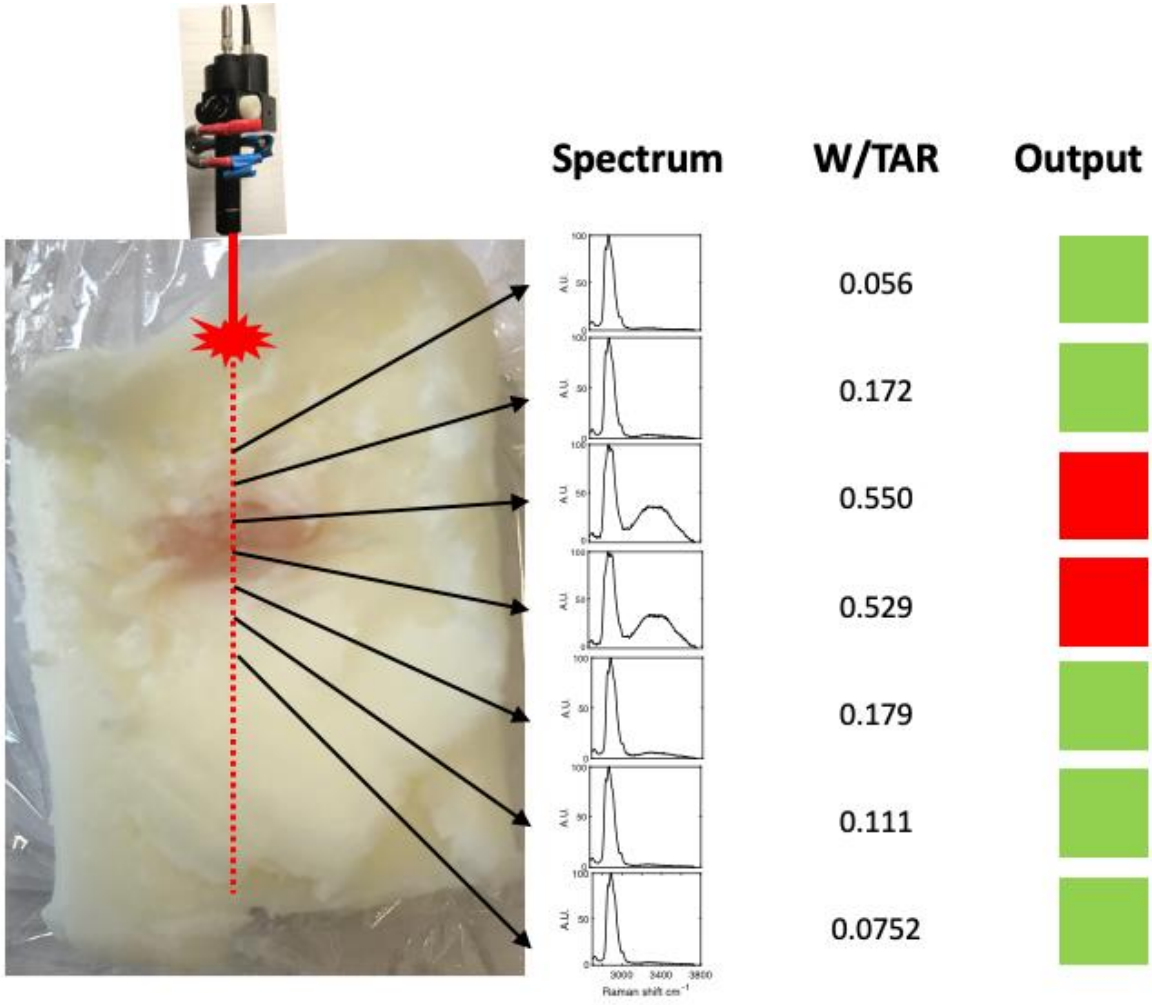


Figure 5-11 demonstrating the concept of a rapid line scan of a lumpectomy phantom

Left-illustration of HH Raman system taking measurements from the lumpectomy phantom, and the line of measurements to be taken (dotted line). Measurements taken at 1 mm step size in a line on the phantom, the corresponding spectrum is displayed, and water/total area ratio (W/TAR) calculated. A visual output is provided based on a W/TAR cut off of 0.5 to differentiate tumour (≥ 0.5; green) from non-tumour (<0.5; red). Raman spectra are single spectrum with 1 second acquisition, pre-processed and normalised to CH stretch region between 2850 – 2950 cm^{-1}

There are limitations to this illustrative experiment. Although the point measurements had suggested that 'tumours' of 0.5 mm^2 could be detected, the 1 mm phantom was used because the pork meat in the 0.5 mm phantom could

not be accurately seen to confirm it was being measured in the dark laser safe room. A manual single line scan directly through an area of known 'tumour' will obviously have an improved detection rate compared to an automated scan through tissue of unknown composition. The spectral differences between pork meat and lard will undoubtedly be easier to identify than the more subtle changes that can be seen between breast tumour and non-tumour tissue.

Despite these limitations, this experiment illustrates the potential in the HH Raman system for providing IMA. These results demonstrate that a small area of 'tumour' (1mm^2) could be easily detected with the HH Raman system in a lumpectomy phantom measured in a line scan, with acquisition times of 1 second. The calculation of the water/total area ratio allows for rapid assessment of the Raman spectra and could give outputs that are easy to interpret by the surgeon to assist intraoperative decision making.

5.4 Discussion

The series of experiments in this chapter demonstrates the technical capabilities and limitations of the HH Raman system, focused on its clinical application to providing IMA in human breast tissue.

5.4.1 Clinical context of findings

The results from the gelatine phantoms of varying water concentrations demonstrate that the HH Raman system can quantify changes in water content in a protein rich environment, and the results are very similar to those obtained with the NP3 system. This suggests that the previous work performed with the NP3 system concerning water content, and the influences on the HWN spectral region, remain valid with the new HH Raman system.

The experiments concerning laser power, acquisition time and probe to specimen distance were designed to assess the influence of changing these parameters on the resultant Raman spectra. The experiments found that Raman spectra with an adequate signal to noise can still be obtained with the laser power as low as 80 mW, which is a laser power utilised in other studies using Raman *in vivo* [223, 363]. Although the water/total area ratio was significantly different to the ratio obtained

with 250 mW power, it would be unlikely to have a detrimental impact on diagnostic ability.

Adequate Raman spectra could be obtained with signal acquisition times as low as 1 second. Even measurement with an acquisition time of 0.5 seconds did not have a significantly different water/total area ratio compared to the longer acquisition times, however visually the signal to noise was suboptimal and would likely cause diagnostic difficulties. These are significantly shorter acquisition times than previous work with InGaAs cameras in biological tissue, and may be due to our pre-processing methods of dark noise subtraction [345, 365].

The probe is best used with the tip in light contact with the specimen. Measurements performed in this way had the least variability in the water/total area ratio. However, the results were not significantly different from when measurements were taken with the probe not in contact with the specimen and spectra could still be obtained when the probe was at a distance of up to 10 mm away from the specimen. This is reassuring that if there is variability in using the probe between users, or between specimens, the results would not be adversely affected, which is a potential pitfall for handheld probes providing IMA [266].

These effects of the change of measurement parameters are important to note, as during clinical measurements they may alter erroneously, or if these parameters needed to be altered for experimental need, the effect this may have on the Raman spectra can be anticipated, and this information can inform the decisions on experimental protocol.

Experiments were performed to understand the subsurface sampling depth of the HH Raman probe. Many other Raman systems developed for IMA have focused primarily on surface techniques, with no ability to sample below the surface [186, 189, 299], though SORS probes have had a subsurface sampling capability of up to 2 mm [185], and 3 mm [265] . The experiments investigating the subsurface sampling depth of the probe demonstrated that with up to 3 mm of gelatine phantom overlying fat, lipid signal still contributed to the Raman spectra. There was still a spectral difference, and significant difference in the water/total area ratio, between lipid alone and those with gelatine overlying the lipid. However, it

suggests that in a ‘worse case scenario’ if there was a small area of protein rich tumour tissue overlying a lot of lipid tissue the signal may be overwhelmed by the lipid contribution and cause diagnostic inaccuracy. More clinically relevant was the finding that even a small amount of lipid (lard) overlying protein rich tissue (pork meat) resulted in no signal from the protein rich tissue being obtained. Translating this to the provision of IMA, if a tumour has a small amount of normal, fatty tissue overlying it, it may go undetected. Currently the guidance is that further excision should only be performed if there is tumour at the resected edge (tumour ‘on ink’) i.e. an exposed tumour with no normal tissue overlying it [56]. Margins of 0-2mm are subject to discussion about re-excision only, and therefore it is probable they would not take place necessarily at the time of the operation anyway. With these current guidelines therefore, the subsurface sampling depth of the HH Raman probe (of <1 mm in areas of lipid) are ideal for the provision of IMA.

Experiments on lumpectomy phantoms (with pork meat mimicking ‘tumour’, and lard ‘non-tumour’ tissue) gave insight into the sampling area of the HH Raman system. Raman spectra obtained from point measurements of the 0.5 mm lumpectomy phantom had some lipid contribution from the surrounding lard, whereas the spectra from the 1 mm phantom did not, suggesting that the sampling area of the probe is between 0.5 – 1 mm^2. The lipid contribution was visually apparent in the spectra from the 0.5 mm phantom which did reduce the water/total area ratio in this phantom compared to the larger ‘tumour’ phantoms (1-5 mm) but the water/total area ratio was still significantly different between lard and all the lumpectomy phantoms (0.5 – 5 mm). This suggests that ‘tumours’ as small as 0.5 mm^2 could be detected using the HH Raman system, and certainly ‘tumours’ of 1mm^2 or greater could be readily detected. This size of tumour detection would compare very favourably with not only other Raman systems [189] but also other technologies [370] aiming to provide IMA.

The line scan of a lumpectomy phantom with a 1 mm^2 ‘tumour’ (pork meat within lard) demonstrates the potential of the HH Raman system and possible future application. The scan demonstrated the potential to detect a small tumour of 1 mm^2 with 1 second acquisition times, with the water/total area ratio calculated simultaneously with a ‘real time’ easily interpretable output to assist clinical

decision making. The line scan shows that the entire resected edge of a small lumpectomy specimen (30 X 30 mm) could be analysed in 15 minutes with this system (assuming 1 mm step size). This is potentially a highly favourable specimen analysis time when compared to other Raman systems [156, 189], and other methods of IMA such as frozen section or cytology [136] and is a clinically relevant time.

5.4.2 Limitations

Many of these experiments compared pork meat to pork fat/lard through the water/total area ratio. Although fat/lard had consistent water/total area ratio, pork meat had more heterogenous findings. This is due to the natural variation in pork meat, but also its water content, as the pork meat may have had different initial water content and it may dehydrate at different rates according to the size it was dissected into (i.e. the difference in surface area between 0.5 mm^2 and 5 mm^2). This led to some variation between water/total area ratios between measurements of pork meat that may not have been expected (e.g. in the lumpectomy specimens the 2 mm and 3 mm phantoms had lower water/total area ratios than the 1 mm and 4 mm phantom). However, this heterogenous tissue, with a variation of water content and water/total area ratio makes it more comparable to human breast tissue than a more experimentally produced phantom with a fixed water content. It should also be noted that these (relatively) minor variations between pork meat tissue did not diminish the large spectral differences between pork meat and pork fat and the general findings and conclusions.

A specific limitation of the subsurface sampling depth experiments was the inability to accurately measure the depth of the material (gelatine phantom or lard) overlying the base tissue (pork fat or meat). This means that the results are estimations of sampling depths only. The lard overlying meat was particularly limiting as no meat signals could be obtained, it would be interesting to see if that would still be true with only microns of lard overlying. Due to these limitations the subsurface sampling depth of the probe could not be accurately quantified, however, clinically relevant deductions regarding the ability to detect tumour beneath normal tissue could still be surmised.

A general limitation of the experiments is that pork meat and fat and lard may not have the same Raman cross section as breast tissue and the results found with pork tissue may not be precisely validated in human breast tissue. Similarly, spectral differences between pork meat and fat are well defined and precise, compared to more heterogenous breast tissue. The ability to differentiate meat and fat in these experiments may not equal the diagnostic ability of the system in breast tissue. However, the aim of these experiments was to provide a close representation of what could be expected with breast tissue and pork tissue is the most suitable phantom for breast tissue, as has previously been demonstrated in this thesis, and in other studies [371].

The line scan measurements performed at 1 second acquisition time show great promise in reducing specimen analysis time, however, it is not just signal acquisition time that dictates this. The entire specimen analysis time would need to include the movement between step measurements of an electronic stage, or the turning of a specimen on a platform to allow measurement of another resected edge. Working to reduce the specimen analysis time will be necessary in the future.

5.4.3 Further work

Further work with the HH Raman probe will be with human breast tissue. With the understanding of the probe capabilities demonstrated in this chapter, measurements on fresh breast tissue can be performed. If it is found e.g. that the laser power needs to be reduced, or an uneven specimen surface results in poor probe contact with the specimen, the results can now be interpreted accordingly. The main emphasis is on using fresh breast tissue to determine the ability of HH Raman system to differentiate between tumour and non-tumour tissue. Once this has been ascertained, more refined diagnostic capabilities such as the minimal detectable tumour size and reducing specimen analysis time can be the focus.

Optimisation of the HH Raman system towards IMA provision can continue alongside clinical studies in human breast tissue. Acquisition times of 1 second are plausible with the HH Raman system immediately, however, with some further work shorter acquisition times look likely to be obtainable. Further work such as reducing the laser induced background of the probe, optimisation of

fluorescent subtracting baselining, further InGaAs camera cooling [339] and a higher quality InGaAs camera with less inherent dark noise [365] are all possible and could lead to reduced signal acquisition times. Introducing a mechanical electronic stage scan, and a sampling procedure to minimise the overall specimen analysis time would all be steps towards translating this technology into the clinical sphere.

5.5 Conclusion

This chapter has presented a new HH Raman system that was constructed to address the recognised limitations of the NP3 system that has been used previously throughout the thesis. A variety of the probe's characteristics have been explored, demonstrating that laser power, signal acquisition time and probe to specimen contact can be changed according to the needs of future experimental design, and the impact on these changes on Raman spectra are demonstrated. The limitations of the probes subsurface sampling depth are now recognised, but this is unlikely to have a negative clinical effect in providing IMA. Ultimately the future potential of the probe is demonstrated, in a pork meat/lard lumpectomy phantom a line scan could be performed with 1 second signal acquisition identifying a 'tumour' of 1 mm^2. Future work identified will be in human breast tissue validating these findings and assessing the diagnostic ability of the probe. This is a robust Raman system that has the potential to measure small tumours in a clinically relevant time and is the next stage in project development towards successfully providing IMA.

CHAPTER 6: Discussion

6.1 Summary of Thesis findings

In the background to this thesis (chapter 1), the primary clinical problem was outlined – a large proportion (17%) of women undergoing BCS in the UK have a second operation for re-excision of margins [91]. This represents a substantial cost to health services and is a source of anxiety for patients [372]. If it was possible to provide accurate Intraoperative Margin Analysis (IMA), this would improve patient outcomes and reduce the cost of treating breast cancer in a resource-stretched NHS. Raman spectroscopy is capable of rapid, non-invasive tissue analysis providing excellent diagnostic accuracy and shows promise for IMA in a number of pathologies, and a detailed review of the current literature is presented in chapter 1. A common limitation for Raman spectroscopy to provide IMA is that the time taken to analyse a specimen can be impractical, due to spectral acquisition time and small sampling volumes. The increased spectral intensity in the HWN region can reduce measurement times compared to the fingerprint region, thus making it an ideal emerging technology that warrants investigation to provide IMA. This aim of the thesis was to investigate the potential of HWN RS to provide IMA using its high sensitivity to detect protein, lipid and particularly water changes, and explore the ability of HWN RS to differentiate tumour from non-tumour in breast tissue.

The first step was to develop a Raman system configuration that was capable of accurately detecting changes in protein, lipid and water content. Chapter 2 demonstrates the evolution of the Raman system. All three Raman systems that were trialled could measure changes in water content. However, I found that surgical pigments (particularly blue dye) that are commonly used in BCS produced overwhelming fluorescence with a 680 nm laser excitation wavelength. Many previous Raman studies either avoid samples with blue dye [299], wash samples or avoid areas with surgical pigments on [174, 189], however, this is impractical for clinical *ex vivo* assessments. A Raman system to provide IMA must be accurate in the presence of surgical pigments. I found that capturing the HWN region, while avoiding the strong fluorescent patent blue dye, required a

785 nm laser with an InGaAs camera – the NP3 system. We are the first group to directly address the issue of blue dye fluorescence when capturing the HWN region and NP3 is the first Raman system that has been shown to capture the HWN region in the presence of blue dye.

In chapter 3 I investigated how the HWN spectra change in response to water content, particularly in biological samples. I demonstrated that the Raman water peak, and changes with water content of a sample, are affected by the microenvironment (whether it was in a protein or lipid rich area) of the water. This means that in a biological sample, such as breast tissue, when measurements are taken from different microenvironments it is difficult to quantify differences in water content with HWN RS. Previous studies have not acknowledged the importance that the microenvionment of water has on the ability of HWN RS to measure changes in water content [174]. These new insights into the influences of the microenvironment on the ability of HWN RS to quantify water content in complex biological systems help to deepen the understanding of the technique within the field. However, the experiments did confirm that changes in water content within a homogenous environment can be quantified accurately, and that HWN spectral analysis using the water/total area ratio can be used to measure differences in water content, and differences in protein or lipid rich areas.

Human breast samples were measured in chapter 4 to investigate if HWN Raman, and NP3 could diagnose breast cancer with a view to IMA. It showed that breast tumour tissue is a protein rich, high water, low fat environment and that non-tumour is a low protein, fat rich environment with a low water content, and, for the first time, it was shown that these HWN features can be used to differentiate tumour from non-tumour breast tissue using HWN RS with excellent accuracy. A number of different spectral analyses were undertaken, with similar diagnostic accuracies – of note – the simplest analysis using the water/total area ratio had a comparable diagnostic accuracy to complex concatenated spectral analysis using PCA fed LDA analysis. No studies have previously compared the diagnostic ability of Raman spectra from the HWN region compared to the FP region taken from the same measurement spot in breast tissue, and we found that the HWN region had a similar diagnostic accuracy, if not marginally better, than the FP region. It was also encouraging that our HWN Raman system was

capable of differentiating DCIS only tissue from non-tumour tissue, which is important in the provision of IMA, and has not been assessed in other studies investigating the HWN region.

Having performed the proof of principle studies, the NP3 system configuration was optimised for clinical use by the development of the HH probe in chapter 5. The technical capabilities of the probe were assessed, and studies in lumpectomy phantoms demonstrated the potential of the system to identify small (<1 mm) tumours in measurement times that would allow analysis of a resected specimen edge in a clinically relevant timeframe (15 minutes).

Ultimately this thesis has shown, for the first time, the excellent diagnostic capability of a Raman system to differentiate tumour from non-tumour using HWN RS and this can be performed in the presence of blue dye. The HH probe and measurement protocol suitable for clinical use has been developed for the next experimental stage to translate these findings towards the provision of IMA. These results need to be considered compared to other methods of IMA, and how to further these thesis findings and Raman systems into a working clinical device for IMA, which is now discussed.

6.2 Clinical applications of thesis findings

The primary aim was to develop a Raman system to provide IMA. However, some of the findings from this thesis have wider implications. In 2018 The Association of Breast Surgery Surgical Gap Analysis Working Group published a report of suggested areas of focus for future research relating to breast surgery. Raman spectroscopy is well suited to a number of these research areas and exploration for the potential applications of the technique is warranted [98]. A brief summary of how some of the findings from this thesis could be exploited for clinical use, or areas worthy of further research identified by the ABS Gap analysis (identified by *) are highlighted.

- *Biochemical understanding of breast cancer*. We demonstrate the differences between tumour and non-tumour tissue, in the differences in protein/lipid, and the substantial differences in water content. These have previously only been demonstrated in rat models, or in a highly controlled

laboratory environment. The results in chapter 4 robustly corroborate these findings in human breast tissue, suggesting these biochemical differences are easily detectable with HWN RS.

- *Biochemical differences in non-tumour tissue*. We found that there is a significant physiological variation in water content in normal, non-tumour specimens related to age and menopausal status. This may be important in understanding the role of Raman spectroscopy in the diagnosis of disease states. For example, breast density (which is closely related to water content) is an independent breast cancer risk factor [280]. This detection of physiological differences with HWN RS could be utilised in screening or risk assessment etc. for personalised medicine to identify patients at an increased risk of breast cancer e.g. those in need of earlier or more intensive screening for breast cancer as a way to reduce screening overdiagnosis (*).
- *Diagnosis / tumour guidance with needle biopsy*. We have shown that a Raman system using a needle biopsy (22G) is clinically effective and can take measurements to diagnose cancer. This demonstrates that HWN RS could be used percutaneously for instant biopsy diagnosis to improve cost-effectiveness of symptomatic assessment (*), tumour localisation guidance (i.e. in replacement of wire guided tumour localisation for impalpable tumours) or local brachytherapy.
- *Translating findings to Deep Raman.* Using the understanding of changes in water content from this thesis has already been translated into furthering understanding of HWN RS in Deep Raman, with a view to providing transcutaneous Raman measurements for the screening program or to complement Mammography [371].
- *Neoadjuvant therapy monitoring(*)*. An exciting prospect for this specific Raman system, would be for monitoring patient tumour response to neoadjuvant chemotherapy. Indications for neoadjuvant therapy are increasing, and yet there is no accurate way of assessing tumour response. Tumour water content changes in response to chemotherapy [289], and HWN RS may provide an accurate method for assessing patient response.

6.3 Comparing findings to other methods of IMA

It is important to consider the clinical implications of the diagnostic accuracy of a system that aims to provide IMA. The aim, as with any diagnostic test, is to have as high a sensitivity and specificity as possible, but in reality, there is a balance between the two. A system that gives a high degree of sensitivity and low specificity would reduce the need for delayed re-excision of margins, but would lead to a high number of false positives, unnecessary breast tissue excision and an undesirable cosmetic impact [373]. A system that has a low sensitivity and high specificity may not identify positive margins and not reduce the delayed re-excision rate sufficiently enough to be of practical use. The sensitivity and specificity of our system presented in this thesis needs to be considered in the context of other methods of IMA for breast conserving surgery.

A number of established techniques have been previously used to provide IMA in breast conserving surgery, however, none have gained widespread adoption – invariably due to logistical difficulties or ultimately poor clinical outcomes. A review of the commonly used techniques is in Chapter one of this thesis, a summary of the sensitivity and specificity of established techniques along with their clinical advantages and disadvantages is displayed in Table 6-1. These results show that radiological techniques of ultrasound and X-ray give very poor sensitivity, and thus do not adequately reduce the need for re-excision of margins. Frozen section and cytology are both pathological techniques that require real-time histopathology/cytology expertise availability. Although this gives them the best diagnostic accuracy, it is also their disadvantage as these are expensive services to run – providing the service is only cost effective when there are unusually high rates of delayed re-excision of over 36% [132]. It may also be that the diagnostic accuracy found in funded research studies from specialist centres may not translate to equitable clinical benefit in a more ‘normal’ setting of a time-pressured, non-academic, routine service. This is the likely reason for these methods not being routinely used, but despite these disadvantages, frozen section and cytology have the best sensitivity and specificity, and this accuracy would have to be matched or improved by any new technology to be widely adopted.

IMA TECHNIQUE	SENS.	SPEC.	ADVANTAGES	DISADVANTAGES
FROZEN SECTION	86%	96%	Established technology and used in other surgical oncology	Expensive as requires histopathologist on stand-by
CYTOLOGY	91%	95%	Proven efficacy	Expensive. Surface technique only - Cannot provide information on margin width. Decline in cytology skills within workforce
ULTRA-SOUND	59%	81%	Cheap and available technology	Operator dependent. Requires surgeon to be trained to operate. Poor accuracy
SPECIMEN RADIOLOGY (X-RAY)	53%	84%	Cheap and available technology	X ray interpretation inaccurate. Poor accuracy.

Table 6-1 Table demonstrating the sensitivity (Sens.) and specificity (spec.) of commonly used Intraoperative Margin Analysis (IMA) techniques

The figures for sensitivity and specificity obtained from the pooled meta-analysis of techniques from reference [136]. Studies were clinical studies where the IMA technique outcome (positive/negative margin) was compared to the final histopathology report.

There are a number of new, emerging technologies that aim to provide IMA. Some have been available for a number of years, have undergone assessment in the clinical environment, and are approved by regulatory bodies.

The MarginProbe ® (DUNE Medical Devices, Alpharetta, Georgia, USA) is a handheld probe that provides IMA on resected specimens using radiofrequency. In a study that assessed the diagnostic accuracy of the probe compared to histopathology diagnosis, the sensitivity was dependent on the size of the tumour – with an overall sensitivity and specificity of 70% [369]. The diagnostic accuracy was dependent on the size of the tumour (large tumours had very good accuracy, however, it is unlikely there are large tumours at the resected edge that would require IMA), and heterogenous tumours affected the specificity. Despite these limitations the device has had a number of clinical trials, stating promising initial results, in one study there was a 57% reduction in re-operation rates, however the ‘control’ arm (that did not undergo IMA with the device) had a re-excision rate

of 30%, which is substantially higher than the UK average of 17% [91], and the re-excision rate was reduced to 14% [374]. Another, subsequent study had similar results – a high re-excision rate in the control arm of 39% with the re-excision rate being reduced to 17% (the current UK average) [370]. Neither of the studies reduced the re-excision rate substantially below the current UK average. However, these results have attracted enough attention for the MarginProbe to be the subject of randomised control trial in the UK, the results of which are awaited [375].

The ClearEdge (LS BioPath, Saratoga, CA, USA) is a handheld probe that uses bioimpedance spectroscopy to analyse tissue and has had more limited clinical evaluation. Preliminary results of the diagnostic accuracy of the probe of breast tissue matching to the histopathology could not be found on a literature search, nor referenced in relevant papers, or on the company website. A single clinical study composed of two phases using the device was performed, where the outcome recorded from the probe analysis of a margin was compared to the final histopathology of the specimen margin, and it gave a sensitivity of 84-87% and specificity of 82-75% [266]. A result of note is that in the patients where the device was used appropriately (which did not always occur), the re-excision rate was reduced from 37% (unusually high) to 17% (the current UK average). This diagnostic accuracy is reasonable, however, with no preliminary data it is difficult to compare against this technology, and the clinical results were acceptable but only from a single study, and so further evidence is required before being able to make judgements on the diagnostic accuracy of the device.

In these two ‘at market’ products that have undergone clinical evaluation it is interesting that when using either devices, the re-excision rate was only ‘reduced’ to the current UK average for re-excision (17%), which is the re-excision rate without the use of these devices. It does suggest that, although useful in those units with an initial high re-excision rate, there is significant room for improvement, and the problem of re-excision in breast conserving surgery is unlikely to be solved by currently available products.

Novel techniques that are currently at the development stage have therefore been used to investigate if they can provide an improved clinical outcome. The

iKnife, a monopolar hand piece that analyses the diathermy plume of smoke, was used for analysis of 260 breast tissue specimens that gave a sensitivity of 90.9% and specificity of 98.8% in differentiating between tumour and non-tumour samples [376]. However, validation of results was by whole specimen final pathological diagnosis rather than specific matching of measurement points to a corresponding histological analysis. Another obvious limitation of this technique for IMA is that it requires the diathermy to go into tumour before giving a result, which is not what a surgeon would be aiming for, and it would also not give feedback on 'close' margins.

Optical techniques have been studied to provide the diagnostic accuracy of differentiating tumour from non-tumour tissues, the limitations of which are discussed in chapter 1. Some techniques have relatively poor diagnostic accuracy such as Terahertz imaging (Accuracy of 75%, sensitivity 86% and specificity 66%) [288] and Diffuse Optical Spectroscopy (sensitivity 81-88% and specificity 79-84%) [283]. Optical Coherence Tomography (OCT) had initially promising results, with work in human breast tissue showing an accuracy of 91.9% with a sensitivity of 98.1% and specificity 82.4% [377] at differentiating tumour from fibroglandular tissue, however, when used to provide IMA in cavity shaves from lumpectomy specimens the technique only identified 5 out of 8 specimens with positive margins, giving it a sensitivity of 62.5% [146]. Image interpretation was performed by clinicians, and may have contributed to the poor accuracy, exemplifying the need for clear outcomes from a clinical device.

A number of groups have investigated Raman spectroscopy to examine the accuracy of differentiating tumour from non-tumour specimens, the results of these studies is displayed in Table 6-2.

ARTICLE	RAMAN SYSTEM	SPECIMENS	SENS	SPEC	LIMITATIONS TO CLINICAL USE
HAKA[186]	Handheld probe, point diagnosis	Freshly excised surgical specimens	83%	93%	Poor positive predictive value of 36% Long acquisition times (30secs), with a small sampling volume
KONG[156]	Micro-spectrometry, Raster scanning	Small tissue sections	95.6%	96.2%	Microscopic diagnosis only – not applicable to large tissue volumes of surgical specimens
SHIPP[189]	Combined Auto-fluorescence/Raman microscope	Freshly excised and fresh frozen surgical specimens	95%	82%	Raman measurements from specimens were not always matched with histopathology Only one resected edge could be analysed in clinically relevant time
WANG[270]	SERS	Small sections from Freshly excised surgical specimens	89%	92%	Specimen processing steps for Nanoparticles result in overall long measurement times Small number of samples (n=5)

Table 6-2 Summary table of diagnostic accuracy of different techniques based on Raman spectroscopy for differentiating normal breast tissue from breast tumour.

SERS – surface enhanced Raman scattering. Sens – sensitivity, Spec - specificity

The aim of reviewing these other techniques and previous Raman work is to establish what diagnostic accuracy is required from our own Raman system for it to be clinical useful, and 'competitive' to other more established techniques. It would seem the general accuracy for the established techniques is low, hence why there is a continued need for methods of IMA. A technique with a sensitivity over 85%, and a specificity of over 90% would be comparable to the techniques of frozen section and cytology. Compared to the emerging techniques, including other work from Raman spectroscopy, a system that could achieve both sensitivity of around 90%-92% and specificity of 86-89% would be highly favourable. The results from chapter 4 suggest that with NP3 probe and using the mean concatenated spectrum from homogenous areas in a PCA fed LDA analysis, the system can achieve a sensitivity of 93.8% and a specificity of 92.7%

at differentiating tumour from non-tumour tissue. This ability to differentiate between tumour and non-tumour tissue exceeds the diagnostic accuracy of the 'at market' probes of Marginprobe and ClearEdge, the optical techniques that are currently under development and previous Raman studies. If this translates into clinical practice this is comparable diagnostic accuracy to the pathological methods of frozen section and cytology, and exceeds the commonly used methods of US and specimen X ray. However, there does have to be caution in the interpretation of how our study data may translate into clinical results due to the study limitations.

6.4 Limitations

There are some limitations to the presented work that require consideration and identify areas for future work for translating HWN RS into a clinical effective technology to provide IMA.

We were unable to quantify the changes in water content between tumour and non-tumour breast tissue. Although this would have been useful to achieve in the study, experimental results demonstrated that this is not possible using HWN RS in different breast tissue types. In chapter 3 it was shown that changes in the water peak in the HWN region was influenced as much by changes in the microenvironment of the water, as they were by changes in water content. This was confirmed in the dehydration experiments in chapter 4, where different water/total area ratio values were obtained for entirely desiccated specimens of tumour and non-tumour samples, showing that the ratio values cannot be comparable between tissue types. This means that we cannot assign a water concentration to the spectral findings in human breast tissue. Although this does not impact on the ability to provide IMA, the inability to quantify changes in water content between different breast tissue limits our assessment of the role of water in diagnosis with HWN RS and ability to compare with water values obtained with other techniques. However, the dehydration experiments in chapter 4, and spectral analysis demonstrates that there are substantial differences in water content between tumour (water rich) and non-tumour (water poor) breast tissue, and that these differences form the basis of tissue differentiation with HWN RS.

One limitation is that the human breast tissue measured was tissue that had been frozen. A previous study suggested that there were no significant differences in the Fingerprint Raman spectrum between fresh or frozen breast tissue[189], however, this may not be true in the HWN region, as frozen samples may have a different water content compared to the 'true' water content of fresh tissue. Water content of the specimen is likely to be affected by the freeze/thaw cycle. However, the samples were immediately snap frozen on collection and measured rapidly on thawing, reducing the time for dehydration.

Small biopsy samples were measured in this study, where the water in the tissue may behave differently to the larger samples of breast tissue that would be measured for IMA. Small biopsy samples are often used in preliminary work, which may have substantial biochemical or logistical differences in their sampling methods compared to in the clinical study of an entire lumpectomy specimen. Although it should be noted that one would expect more rapid dehydration from smaller samples (due to the increased surface area), and so the differences in water content between tissue types may actually be underestimated in smaller samples, so it is unlikely these changes in water content would not be seen in larger specimens of fresh tissue.

The measured samples are relatively homogenous samples of tumour tissue (taken from the centre of the tumour) and non-tumour tissue (taken far from the tumour). It may be that the stark difference in water content between the two tissue types, in reality, is more of a gradual change in water content from high (tumour) to low (normal), as was found between 0 – 4 mm from the tumour margin in oral cavity squamous cell carcinoma with HWN RS [228]. A study using DRS suggested that the water content at the border of the tumour may differ from the centre of the tumour and normal breast tissue (>1 cm from the tumour edge) [293]. This gradual change in water content may assist in the identification of tumour margins and 'close' margins, and would benefit from further studies to investigate if there is a gradual change in water content from the tumour edge. The thesis has demonstrated that the difference in water content (as measured by water/total area ratio) between pure tumour and pure normal tissue is substantial, and if there is a gradual change this would very likely be measurable with our Raman system.

The sample size of 96 in the study is a reasonable number, however, it is too small to draw definitive conclusions regarding potential differences between pathological sub-types. There are a number of differences between breast cancers such as receptor status or grade. Our analysis suggested that neither the fingerprint nor HWN region identified significant differences between pathological sub-types, but this may be due to the small sample size, and the study was under-powered to detect subtle differences.

There is some suggestion from our data that lobular carcinoma may prove to be a diagnostic challenge, as it has a lower water/total area ratio than ductal carcinoma, and the Raman micro-spectroscopy studies showed that Raman spectra of lobular carcinoma has similarities to normal breast tissue. Further work is needed, with larger numbers of lobular carcinoma to investigate this. It may be that additional information from the fingerprint spectrum may aid in refining the diagnosis.

There was an insufficient number of DCIS only tumour samples to be confident that the Raman system can differentiate between DCIS only and normal tissue. The ability to detect DCIS only is essential for IMA, as DCIS accounts for a large number of positive margins that require re-excision. The results suggest that the Raman system is capable of diagnosis, and are encouraging, but further work with a greater number of DCIS only samples is required.

This study has provided the proof of principle that was necessary to demonstrate that HWN RS can differentiate between tumour and non-tumour tissue, however the NP3 system is a needle probe, which is not a suitable delivery system for providing IMA as it has a small sampling volume. The time taken for each individual spectral measurement (25 seconds) is also not suitable for clinical use. Ultimately it would take too long to analyse an entire BCS specimen using the delivery system and the acquisition times used in the study in chapter 4.

Many of the limitations of the NP3 system have been addressed in the development of the HH Raman system that is presented in chapter 5. There are still elements that will require further optimisation and can be addressed with

further work developing the delivery system capable of providing IMA in a clinically relevant time. The measurement acquisition times will need to reduce to the minimal amount of time necessary to achieve adequate signal to noise for diagnosis. This study has demonstrated that the spectral differences between tissue types, particularly in the HWN region, are stark and so it is likely this will be achievable as even with a substantial reduction in signal to noise the differences would still be detectable and measurable, particularly if the water/total area ratio was used for diagnosis. In Chapter 5, adequate spectra could be obtained with acquisition times as little as 0.5 seconds, and with some optimisation (such as improved hardware and optimised spectra processing) this could be reduced even further.

Although this study has demonstrated our Raman system has a good diagnostic accuracy differentiating between tumour and not tumour when matched to histopathology, this is different to reducing the rate of re-excision which is the ultimate aim. The number of steps between tissue differentiation and providing IMA are multiple, and the 'diagnostic accuracy' can get reduced at every stage. Initially promising results in laboratory conditions can be disappointing in a clinical setting, as was the case with Diffuse Optical Spectroscopy [148]. Logistical difficulties can also intercede, such as in the ClearEdge study were surgeons misinterpreted the probe results [266]. However, the opposite can also be true such as with the MarginProbe, where initial assessment of tissue differentiation seemed poor, but the clinical results at reducing re-excision rate is more impressive [370].

CHAPTER 7: Further Work and Conclusion

7.1 Further Work

7.1.1 Validation of diagnostic findings in fresh breast tissue

The immediate next phase for the project is to validate the thesis findings in freshly excised breast tissue. A limitation of the breast tissue study is that the specimens were frozen and thawed, which may affect the water content of the specimens, or there may be other biochemical differences between frozen and fresh specimens. Previous studies suggest that there are minimal differences in the Raman spectra between fresh and frozen breast tissue [189], but our findings still require validation in fresh tissue.

The aims would be to confirm and validate the diagnostic ability of HWN RS to differentiate between tumour and non-tumour breast tissue in fresh specimens using the HH Raman system. It would also give us insight into the differences in water content between tissue types, and allow exploration of the concept that there may be a gradual change in water content from the centre of the tumour to the tumour edge and towards normal tissue. We have already gained ethical and local approval for this study (REC ref: 18/NW/0366, IRAS ID: 210732), using freshly excised breast tissue of mastectomy and lumpectomy specimens that will be sliced in half to expose the tumour surface, allowing Raman measurements to be taken from tumour and non-tumour tissue. A diagram will indicate to the pathologist where measurements have been obtained, and this will allow for a block to be taken for histopathological diagnosis to allow pathology matching with the Raman measurements. Initial ethical approval is for 50 patients to be recruited for the study.

Once the initial phase is completed, and if the thesis diagnostic findings are confirmed in fresh tissue, we may need to extend the study. This may be necessary to increase sample size to ensure all pathology subtypes are encountered during analysis, especially adequate numbers of DCIS and lobular carcinoma, as these are potential diagnostic challenges. There should also be some histological matching for normal tissue, to investigate the Raman spectra of benign abnormalities, as this may affect diagnostic accuracy.

7.1.2 Probe optimisation

Alongside validation of diagnostic accuracy of the HH Raman system in differentiating tumour from non-tumour tissue, probe optimisation is required. This is to develop a probe that is capable of analysing a lumpectomy in a clinically relevant time frame (<30 minutes).

Two components of this optimisation are to reduce the signal acquisition times for each measurement, or to reduce the number of measurements required to accurately analyse a specimen. The signal acquisition time needs further investigation to discover what is the shortest signal acquisition time possible that still gives a diagnostic Raman spectrum, this can be done by testing a number of signal acquisition times as part of the fresh breast tissue experiments. The main limitation to reducing the signal acquisition time currently is the high degree of detector noise from the InGaAs camera, which is a known limitation of this type of camera. However, technology is improving, and there are recent reports of novel deep cooled InGaAs cameras that have a lower noise characteristics, comparable to those of CCD cameras [365]. Assessment of the detector noise characteristics of different InGaAs cameras may be necessary to identify the optimal hardware for reducing signal acquisition times.

Another method to reduce signal acquisition times is to reduce the number of spectra required for specimen assessment. If the beam area of the probe was increased to increase the probe's 'field of view', this would reduce the number of measurements required to measure the entire specimen. Further work to assess how much the probe area can be increased without a reduction in diagnostic accuracy is required.

We have developed a handheld probe for the initial phases of the project, and it may be that this is the most suitable device for the clinical environment. The margin of concern to the surgeon can be easily analysed using a clinical probe, and the fact that the other 'at market' devices for IMA are also handheld probes (ClearEdge and MarginProbe) suggests that this might be the most effective way of IMA. However, other devices should be considered, whether that is an automatic device that moves the specimen underneath the already designed

probe (such as an automated stage with rapid raster scanning for analysis of a single margin), or a more complex device allowing all margins to be analysed such as the Marginbot [265]. A recently reported Raman system using a 2 axis galvo-mirror system and a microscope with a large field of view may provide rapid analysis, and our configuration could be incorporated into a similar system [299]. These possibilities should be explored to determine the optimal device for delivering IMA in clinically relevant times.

7.1.3 Data analysis optimisation

The results in this thesis from the analysis of frozen breast tissue suggests that the diagnostic ability from the water/total area ratio and PCA fed LDA analysis of the HWN or FP region and the concatenated spectra were similar. The planned fresh tissue studies will help to clarify this further, and assist in deciding the exact diagnostic route that should be used for IMA; a suggested algorithm based on the thesis results is presented in Chapter 4, but this may alter depending on the results of this further work.

Data analysis needs to be as rapid as possible, to reduce the time taken for specimen analysis. It also needs a simple output for each of the analysed specimen resected edges that is easily interpretable to the surgeon, to allow them to make a clinical decision to excise further tissue or not. A suggested simplified algorithm based just on the water/total area ratio is presented in chapter 5.

7.1.4 Clinical trial

Once these phases of optimisation have been successfully completed, we could be confident that we have a Raman system that can differentiate between tumour and non-tumour tissue accurately, can analyse a lumpectomy specimen in clinically relevant times and the data can be presented in a clinically meaningful way to a surgeon. At this stage a clinical study assessing the ability of the final Raman system to provide IMA is necessary, along with commercialisation of the probe. This will involve recruiting patients undergoing lumpectomy and analysing the margins of specimens and determining the sensitivity and specificity of the system of detecting positive margins compared to the final histopathology report. The next phase is a randomised controlled trial with one arm using the IMA system output to influence the surgeons' clinical decision to resect and the other

arm being surgeons continuing their usual methods of IMA. The measurable outcome will be if the Raman IMA system reduces the re-excision rate with minimal unnecessary breast volume loss.

7.2 Conclusion

This thesis presents a Raman system, and method of spectral analysis, capable of quantifying changes in water content in a number of different environments using HWN RS. This is unaffected by the presence of commonly used surgical pigments. I explored the optical influences on Raman spectra in the HWN region, and concluded the interpretation of the water peak is dependent on its microenvironment. This formed the basis for a study using human breast tissue which demonstrated, for the first time, that HWN RS has an excellent diagnostic accuracy in differentiating tumour from non- tumour specimens. These findings could be useful in a number of novel ways in future clinical applications, but providing IMA is the immediate focus for further work and development. Our study results compare favourably, if not better, than current and emerging methods for IMA.

Further work must focus on assessing the diagnostic ability of HWN RS in larger volumes of freshly excised breast tissue in a large number of patients, and in the future, translating this proof of principle into the operating theatre to improve patient outcomes in BCS and reduce the need for surgical re-excision.

This thesis has demonstrated that high wavenumber Raman Spectroscopy has the potential to provide intraoperative margin analysis in the provision of breast conserving surgery.

Appendices

Appendix A Presentations and publications arising from this thesis

Presentations

Hubbard, Thomas, Douglas Ferguson, Angela Shore, and Nick Stone. Raman spectroscopy for intraoperative margin analysis in breast conserving surgery. Poster presentation (poster 122) accepted for presentation at the Association of Breast Surgeons Conference 2020 Bournemouth (cancelled due to Covid- 19)

Hubbard, Thomas, Douglas Ferguson, Angela Shore, and Nick Stone. Raman spectroscopy for intraoperative margin analysis in breast conserving surgery. Invited speaker oral presentation (Abstract number - RAM-07.1.) at the SciX conference, Sparks, Nevada, October 2020

Published Abstract

Hubbard, Thomas, Douglas Ferguson, Angela Shore, and Nick Stone. "P122: Raman spectroscopy for intraoperative margin analysis in breast conserving surgery." *European Journal of Surgical Oncology* 46, no. 6 (2020): e43.

Publications

Ghita, Adrian, **Thomas Hubbard**, Pavel Matousek, and Nicholas Stone. "Non-invasive Detection of Differential Water Content inside Biological Samples using Deep Raman Spectroscopy." *Analytical Chemistry* (2020).

Hubbard, Thomas James Eliot, Angela Shore, and Nicholas Stone. "Raman spectroscopy for rapid intra-operative margin analysis of surgically excised tumour specimens." *Analyst* 144, no. 22 (2019): 6479-6496.

Appendix B Copyright for images from other sources

Chapter 1

ELSEVIER LICENSE TERMS AND CONDITIONS
Feb 26, 2020
This Agreement between Thomas Hubbard ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.
License Number 4776520637627
License date Feb 26, 2020
Licensed Content Publisher Elsevier
Licensed Content Publication Elsevier Books
Licensed Content Title Gray's Anatomy for Students
Licensed Content Author Richard L. Drake,A. Wayne Vogl,Adam W.M. Mitchell
Licensed Content Date Jan 1, 2020
Licensed Content Pages 133
Start Page 123
End Page 247.e8
Type of Use reuse in a thesis/dissertation
Portion figures/tables/illustrations
Number of figures/tables/illustrations 2
Format both print and electronic
Are you the author of this Elsevier chapter? No
Will you be translating? No
Title Using Raman Spectroscopy for intraoperative margin analysis in Breast Conserving Surgery
Institution name University of Exeter
Expected presentation date Oct 2020
Portions Figure 3.9, cross section of breast Figure 3.16, diagram of Breast
Requestor Location
Thomas Hubbard
United Kingdom Attn: Thomas Hubbard
Publisher Tax ID GB 494 6272 12

This Agreement between Thomas Hubbard ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.
License Number 4777111034911
License date Feb 27, 2020
Licensed Content Publisher: Springer Nature
Licensed Content Publication: Theoretical Chemistry Accounts
Licensed Content Title: High-wavenumber FT-Raman spectroscopy for in vivo and ex vivo measurements of breast cancer
Licensed Content Author A. F. García-Flores et al
Licensed Content Date Apr 6, 2011
Type of Use Thesis/Dissertation
Requestor type academic/university or research institute
Format print and electronic
Portion figures/tables/illustrations
Number of figures/tables/illustrations 1
Will you be translating? no
https://s100.copyright.com/AppDispatchServlet 1/5
27/02/2020 RightsLink Printable License
Circulation/distribution 1 - 29
Author of this Springer Nature content no
Title Using Raman Spectroscopy for intraoperative margin analysis in Breast Conserving Surgery
Institution name University of Exeter
Expected date presentation : Oct 2020
Portions
Figure 1.C. Comparison of mean spectra of normal and cancer tissue in a selected HF region for the three different ways measured. c normal (EVNT) and cancer (EVCT) ex vivo biopsy
Requestor Location
Thomas Hubbard
United Kingdom Attn: Thomas Hubbard
Total 0.00 GBP

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE TERMS AND CONDITIONS
Jul 14, 2020
This Agreement between Thomas Hubbard ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") consists of your license details and

the terms and conditions provided by The American Association for the Advancement of Science and Copyright Clearance Center.
License Number 4867591033406
License date Jul 14, 2020
Licensed Content Publisher: The American Association for the Advancement of Science
Licensed Content Publication: Science Translational Medicine
Licensed Content Title : Intraoperative brain cancer detection with Raman spectroscopy in humans
Licensed Content Author : Michael Jermyn,Kelvin Mok,Jeanne Mercier,Joannie Desroches,Julien Pichette,Karl Saint-Arnaud,Liane Bernstein,Marie-Christine Guiot,Kevin Petrecca,Frederic Leblond
Licensed Content Date : Feb 11, 2015
Licensed Content Volume : 7
Licensed Content Issue : 274
Volume : 7
Issue number : 274
Type of Use : Thesis / Dissertation
Requestor type : Scientist/individual at a research institution
Format Portion : Print and electronic Figure
Number of figures/tables : 1
Title : Using Raman Spectroscopy for intraoperative margin analysis in Breast Conserving Surgery
Institution name : University of Exeter
Expected presentation date : October 2020
Portions : Figure 1.B
Requestor Location : Thomas Hubbard United Kingdom
Terms and Conditions
American Association for the Advancement of Science TERMS AND CONDITIONS
Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission, to republish the AAAS material identified above in your work identified above,

Chapter 2

SPRINGER NATURE LICENSE TERMS AND CONDITIONS
Apr 06, 2020
This Agreement between Thomas Hubbard ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.
License Number 4803141244544
License date Apr 06, 2020
Licensed Content Publisher Springer Nature
Licensed Content Publication Analytical and Bioanalytical Chemistry
Licensed Content Title Characterisation of a fibre optic Raman probe within a hypodermic needle
Licensed Content Author Ingeborg E. Iping Petterson et al
Licensed Content Date Sep 28, 2015
Type of Use Thesis/Dissertation
Requestor type academic/university or research institute
Format print and electronic
Portion figures/tables/illustrations
Number of figures/tables/illustrations 1
Will you be translating? no
Circulation/distribution 1 - 29
Author of this Springer Nature content no
Title Using Raman Spectroscopy for intraoperative margin analysis in Breast Conserving Surgery
Institution name University of Exeter
Expected presentation date Oct 2020
Portions Figure 2
Requestor Location
Thomas Hubbard , United Kingdom

Feb 27, 2020
Royal Society of Chemistry - License Terms and Conditions
This is a License Agreement between Thomas Hubbard ("You") and Royal Society of Chemistry ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Royal Society of Chemistry, and the CCC terms and conditions.
All payments must be made in full to CCC.
Order Date 06-Apr-2020 **Order license ID** 1026927-1 **ISSN** 1364-5528
Type of Use Republish in a thesis/dissertation
Publisher Royal Society of Chemistry
Portion Chart/graph/table/ gure
Publication Title The analyst online
Author/Editor Society of Public Analysts
(Great Britain), Chemical Society (Great Britain), Society for Analytical Chemistry, Society of Public Analysts

(Great Britain), Royal Society of Chemistry (Great Britain)
Date 01/01/1876
Language English
Country United Kingdom of Great Britain and Northern Ireland
Rightsholder Royal Society of Chemistry
Publication Type e-Journal
URL http://www.rsc.org/is/jour nals/current/analyst/anlco n.htm
Portion Type Chart/graph/table/ gure
Number of charts / graphs / tables / gures requested
Distribution U.K. and Commonwealth
1 (excluding Canada)
Translation Original language of publication
Format (select all that : Print, Electronic
Copies for the disabled? No **apply)**
Minor editing privileges? No
Who will republish the content: Academic institution
Duration of Use Life of current edition
Lifetime Unit Quantity Up to 499
Rights Requested Main product
Currency GBP
Title Using Raman Spectroscopy for intraoperative margin analysis in Breast Conserving Surgery
Instructor name Thomas Hubbard
Institution name University of Exeter
Expected presentation date : 2020-10-01
Order reference number N/A
The requesting person / organization to appear on the license: Thomas Hubbard
REUSE CONTENT DETAILS
Title, description or numeric reference of the portion(s) : Figure 3
Editor of portion(s) : N/A
Volume of serial or monograph : N/A
Page or page range of portion :6055
Title of the article/chapter the portion is from :
Dual excitation wavelength system for combined ngerprint and high wavenumber Raman spectroscopy
Author of portion(s)
Society of Public Analysts (Great Britain); Chemical Society (Great Britain); Society for Analytical Chemistry; Society of Public Analysts (Great Britain); Royal Society of Chemistry (Great Britain)
Publication date of portion : 2018-10-27

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424.
2. Office of National Statistics (UK). Cancer registration statistics, England:2016 [Internet]. UK Government; 2016 [Accessed 21/01/2019]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/final2016.
3. NHS England. Cancer Strategy Implementation Plan [Internet]. NHS England; 2018 [Accessed 03/12/2018]. Available from: https://www.england.nhs.uk/cancer/strategy/.
4. Igene H. Global health inequalities and breast cancer: an impending public health problem for developing countries. *Breast J*. 2008;14:428-434.
5. Dixon JM. ABC of breast diseases: John Wiley & Sons; 2012.
6. Dixon JM. The Breast. In: Garden OJ PR, editor. Principles and practice of Surgery. 1. 7th ed. Edinburgh: Elsevier; 2018. p. 326 - 349.
7. Kliewer EV, Smith KR. Breast cancer mortality among immigrants in Australia and Canada. *J Natl Cancer Inst*. 1995;87:1154-1161.
8. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *Lancet*. 1996;347:1713-1727.
9. Kispert S, McHowat J. Recent insights into cigarette smoking as a lifestyle risk factor for breast cancer. *Breast Cancer*. 2017;9:127 - 132.
10. Begg CB, Haile RW, Borg Å, Malone KE, Concannon P, Thomas DC*, et al.* Variation of breast cancer risk among BRCA1/2 carriers. *JAMA*. 2008;299:194-201.
11. Newman B, Austin MA, Lee M, King M-C. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A*. 1988;85:3044-3048.

12. National Institute for Health and Care Excellence (NICE). Familial breast cancer: classification, care and managing breast cancer and related risks in people with a family history of breast cancer [CG164] [Internet]. NICE; 2017 [Accessed 29/01/2019]. Available from: https://www.nice.org.uk/guidance/cg164/chapter/Recommendations#clinical-significance-of-a-family-history-of-breast-cancer.

13. Lawrence G, Wallis M, Allgood P, Nagtegaal ID, Warwick J, Cafferty FH, *et al.* Population estimates of survival in women with screen-detected and symptomatic breast cancer taking account of lead time and length bias. *Breast Cancer Res Treat*. 2009;116:179-185.

14. Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study. *Lancet*. 2015;385:1206-1218.

15. Cancer Research UK (CRUK). Cancer Research UK: Breast Cancer survival statistics [Internet]. CRUK; 2018 [Accessed 12/12/2018]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/survival#heading-Two.

16. Campbell JD, Ramsey SD. The costs of treating breast cancer in the US. *Pharmacoeconomics*. 2009;27:199-209.

17. Drake RL VA, Mitchell AWM. Gray's Anatomy for Students. 4th ed. Elsevier: Philadelphia; 2020. 1103 p.

18. Cooper AP. On the Anatomy of the Breast: Longman; 1840.

19. Tobon H, Salazar H. Ultrastructure of the human mammary gland. II. Postpartum lactogenesis. *J Clin Endocrinol Metab*. 1975;40:834-844.

20. Ramsay D, Kent J, Hartmann R, Hartmann P. Anatomy of the lactating human breast redefined with ultrasound imaging. *J Anat*. 2005;206:525-534.

21. Brisken C. Hormonal control of alveolar development and its implications for breast carcinogenesis. *J Mammary Gland Biol Neoplasia*. 2002;7:39-48.

22. Anderson E DC, Norman J, Payne S. The Reproductive System. In: Douglas G NF, Robertson C, editor. Macleod's Clinical Examination. 13 th ed. Edinburgh: Elsevier; 2013. p. 212 - 218.

23. Kumar V AA, Aster J. Neoplasia. In: Kumar V AA, Aster J, editor. Robbins Basic Pathology. ninth ed. Philadelphia: Elsevier; 2013. p. 161-163.

24. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C*, et al.* Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer*. 2015;51:2296-2303.

25. Kumar V AA, Aster J. Female Genital System and Breast. In: Kumar V AA, Aster J, editor. Robbins Basic Pathology. Phildadelphia: Elsevier; 2013. p. 704-714.

26. Henderson I. Breast Cancer: Fundamentals of Evidence Based Disease Management. First ed. Oxford: Oxford University Press; 2015.

27. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer*. 2005;103:2481-2484.

28. Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson IC. Incidence of and treatment for ductal carcinoma in situ of the breast. *JAMA*. 1996;275:913-918.

29. Fox MS. On the diagnosis and treatment of breast cancer. *JAMA*. 1979;241:489-494.

30. Buzdar AU, Ibrahim NK, Francis D, Booser DJ, Thomas ES, Theriault RL*, et al.* Significantly higher pathologic complete remission rate after neoadjuvant therapy with trastuzumab, paclitaxel, and epirubicin chemotherapy: results of a randomized trial in human epidermal growth factor receptor 2–positive operable breast cancer. *J Clin Oncol*. 2005;23:3676-3685.

31. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H*, et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001;98:10869-10874.

32. Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A*, et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*. 2003;100:8418-8423.

33. Koscielny S, Tubiana M, Le M, Valleron A, Mouriesse H, Contesso G*, et al.* Breast cancer: relationship between the size of the primary tumour and the probability of metastatic dissemination. *Br J Cancer*. 1984;49:709.

34. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer*. 1989;63:181-187.

35. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F*, et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet*. 2005;365:671-679.

36. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*. 1991;19:403-410.

37. Sobin L GM, Wittekind C. TNM Classification of Malignant Tumours. 7th ed. L S, editor. UK: Wiley-Blackwell; 2009.

38. Haybittle J, Blamey R, Elston C, Johnson J, Doyle P, Campbell F*, et al.* A prognostic index in primary breast cancer. *Br J Cancer*. 1982;45:361.

39. Todd J, Dowle C, Williams M, Elston C, Ellis I, Hinton C*, et al.* Confirmation of a prognostic index in primary breast cancer. *Br J Cancer*. 1987;56:489.

40. University of Cambridge. Predict! [Internet]. NHS; 2019 [Accessed 22/11/2019]. Available from: https://breast.predict.nhs.uk/.

41. National Institution for Health and Care Excellence (NICE). Suspected cancer: recognition and referral [NG12] [Internet]. NICE; 2015 [Accessed 29/01/2019]. Available from: https://www.nice.org.uk/guidance/ng12.

42. Richards M, Smith P, Ramirez A, Fentiman I, Rubens R. The influence on survival of delay in the presentation and treatment of symptomatic breast cancer. *Br J Cancer*. 1999;79:858.

43. Walker S, Hyde C, Hamilton W. Risk of breast cancer in symptomatic women in primary care: a case–control study using electronic records. *Br J Gen Pract*. 2014;64:e788-e793.

44. Public Health England. NHS Breast Screening Programme Clinical guidance for breast cancer screening assessment. In: Public Health England, editor. online2016.

45. National Institute for Health and Care Excellence (NICE). Clinical Knowledge Summary: Breast screening [Internet]. NICE; 2017 [Accessed 29/01/2019]. Available from: https://cks.nice.org.uk/breast-screening#!scenario.

46. Tabar L, Gad A, Holmberg L, Ljungquist U, Group KCP, Fagerberg C*, et al.* Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet*. 1985;325:829-832.

47. Laudicella M, Walsh B, Burns E, Smith PC. Cost of care for cancer patients in England: evidence from population-based patient-level data. *Br J Cancer*. 2016;114:1286-1292.

48. Marmot M, Altman D, Cameron D, Dewar J, Thompson S, Wilcox M. The benefits and harms of breast cancer screening: an independent review. *Br J Cancer*. 2013;108:2205.

49. Beau AB, Lynge E, Njor SH, Vejborg I, Lophaven SN. Benefit-to-harm ratio of the Danish breast cancer screening programme. *Int J Cancer*. 2017;141:512-518.

50. Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ*. 2009;339:b2587.

51. Bond M, Pavey T, Welch K, Cooper C, Garside R, Dean S*, et al.* Psychological consequences of false-positive screening mammograms in the UK. *Evid Based Med*. 2013;18:54-61.

52. de Gelder R, van As E, Tilanus-Linthorst MM, Bartels CC, Boer R, Draisma G*, et al.* Breast cancer screening: evidence for false reassurance? *Int J Cancer*. 2008;123:680-686.

53. Whelehan P, Evans A, Wells M, Macgillivray S. The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. *Breast (Edinburgh, Scotland)*. 2013;22:389-394.

54. Office of National Statistics (UK). Breast Screening Programme 2016-17 [Internet]. NHS Digital,; 2018 [Accessed 29/09/2020]. Available from: https://files.digital.nhs.uk/pdf/m/f/breast_screening_programme__england__2016-17_-_report__v2.pdf.

55. Biller-Andorno N, Jüni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med*. 2014;370:1965-1967.

56. National Institute for Health and Care Excellence (NICE). Early and locally advanced breast cancer: diagnosis and management [NG101] [Internet]. NICE; 2018 [Accessed 24/01/2019]. Available from: https://www.nice.org.uk/guidance/ng101.

57. Flobbe K, Bosch AM, Kessels AG, Beets GL, Nelemans PJ, von Meyenfeldt MF*, et al.* The additional diagnostic value of ultrasonography in the diagnosis of breast cancer. *Arch Intern Med*. 2003;163:1194-1199.

58. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S*, et al.* Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med*. 2005;353:1773-1783.

59. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG*, et al.* Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA*. 2012;307:1394-1404.

60. Berner A, Davidson B, Sigstad E, Risberg B. Fine-needle aspiration cytology vs. core biopsy in the diagnosis of breast lesions. *Diagn Cytopathol*. 2003;29:344-348.

61. Bundred SM, Maxwell AJ, Morris J, Lim YY, Harake MJ, Whiteside S*, et al.* Randomized controlled trial of stereotactic 11-G vacuum-assisted core biopsy for the diagnosis and management of mammographic microcalcification. *Br J Radiol*. 2016;89:20150504.

62. Burrell HC, Evans AJ, Wilson ARM, Pinder SE. False-negative Breast Screening Assessment. What Lessons Can We Learn? *Clin Radiol*. 2001;56:385-388.

63. Khoury T, Sait S, Hwang H, Chandrasekhar R, Wilding G, Tan D*, et al.* Delay to formalin fixation effect on breast biomarkers. *Mod Pathol*. 2009;22:1457.

64. The Royal College of Pathologists. Guidelines for non-operative diagnostic procedures and reporting in breast cancer screening. In: Non-operative Diagnosis Working Group of the UK National Coordinating Committee for Breast Pathology, editor.: RCPath; 2016.

65. Upile T, Fisher C, Jerjes W, El Maaytah M, Searle A, Archer D*, et al.* The uncertainty of the surgical margin in the treatment of head and neck cancer. *Oral Oncol*. 2007;43:321-326.

66. Makary MA, Epstein J, Pronovost PJ, Millman EA, Hartmann EC, Freischlag JA. Surgical specimen identification errors: a new measure of quality in surgical care. *Surgery*. 2007;141:450-455.

67. Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson AN*, et al.* Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*. 2015;313:1122-1132.

68. Frierson JHF, Wolber RA, Berean KW, Franquemont DW, Gaffey MJ, Boyd JC*, et al.* Interobserver Reproducibility of the Nottingham Modification of the Bloom and Richardson Histologic Grading Scheme for Infiltrating Ductal Carcinoma. *Am J Clin Pathol*. 1995;103:195-198.

69. Robbins P, Pinder S, De Klerk N, Dawkins H, Harvey J, Sterrett G*, et al.* Histological grading of breast carcinomas: a study of interobserver agreement. *Hum Pathol*. 1995;26:873-879.

70. Mays N. Reducing unwarranted variations in healthcare in the English NHS. *BMJ*. 2011;342:d1849.

71. Fleissig A, Jenkins V, Catt S, Fallowfield L. Multidisciplinary teams in cancer care: are they effective in the UK? *Lancet Oncol*. 2006;7:935-943.

72. Whelan J, Griffith C, Archer T. Breast cancer multi-disciplinary teams in England: much achieved but still more to be done. *Breast J*. 2006;15:119-122.

73. Saini K, Taylor C, Ramirez A-J, Palmieri C, Gunnarsson U, Schmoll H-J*, et al.* Role of the multidisciplinary team in breast cancer management: results from a large international survey involving 39 countries. *Ann Oncol*. 2011;23:853-859.

74. Lamb BW, Brown KF, Nagpal K, Vincent C, Green JSA, Sevdalis N. Quality of Care Management Decisions by Multidisciplinary Cancer Teams: A Systematic Review. *Ann Surg Oncol*. 2011;18:2116-2125.

75. Hong NJL, Wright FC, Gagliardi AR, Paszat LF. Examining the potential relationship between multidisciplinary cancer care and patient survival: an international literature review. *J Surg Oncol*. 2010;102:125-134.

76. Kesson EM, Allardice GM, George WD, Burns HJ, Morrison DS. Effects of multidisciplinary team working on breast cancer survival: retrospective, comparative, interventional cohort study of 13 722 women. *BMJ*. 2012;344:e2718.

77. Li CI, Daling JR, Malone KE. Incidence of invasive breast cancer by hormone receptor status from 1992 to 1998. *J Clin Oncol*. 2003;21:28-34.

78. Francis PA, Regan MM, Fleming GF, Láng I, Ciruelos E, Bellet M*, et al.* Adjuvant ovarian suppression in premenopausal breast cancer. *N Engl J Med*. 2015;372:436-446.

79. Early Breast Cancer Trialists' Collaborative Group. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet*. 2011;378:771-784.

80. Cuzick J, Sestak I, Baum M, Buzdar A, Howell A, Dowsett M*, et al.* Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. *Lancet Oncol*. 2010;11:1135-1141.

81. Regan MM, Neven P, Giobbie-Hurder A, Goldhirsch A, Ejlertsen B, Mauriac L*, et al.* Assessment of letrozole and tamoxifen alone and in sequence for postmenopausal women with steroid hormone receptor-positive breast cancer: the BIG 1-98 randomised clinical trial at 8· 1 years median follow-up. *Lancet Oncol*. 2011;12:1101-1108.

82. Early Breast Cancer Trialists' Collaborative Group. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *Lancet*. 2012;379:432-444.

83. Gianni L, Dafni U, Gelber RD, Azambuja E, Muehlbauer S, Goldhirsch A*, et al.* Treatment with trastuzumab for 1 year after adjuvant chemotherapy in patients with HER2-positive early breast cancer: a 4-year follow-up of a randomised controlled trial. *Lancet Oncol*. 2011;12:236-244.

84. Fisher B, Bauer M, Margolese R, Poisson R, Pilch Y, Redmond C*, et al.* Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. *N Engl J Med*. 1985;312:665-673.

85. Overgaard M, Jensen M-B, Overgaard J, Hansen PS, Rose C, Andersson M*, et al.* Postoperative radiotherapy in high-risk postmenopausal breast-cancer patients given adjuvant tamoxifen: Danish Breast Cancer Cooperative Group DBCG 82c randomised trial. *Lancet*. 1999;353:1641-1648.

86. Potter MA. Principles of the surgical management of cancer. In: O. James Garden RWP, editor. Principles and Practice of Surgery. 7th ed. Edinburgh: Elsevier; 2018. p. 86-97.

87. Morris AD, Morris RD, Wilson JF, White J, Steinberg S, Okunieff P*, et al.* Breast-conserving therapy vs mastectomy in early-stage breast cancer: a meta-analysis of 10-year survival. *Cancer J Sci Am*. 1997;3:6-12.

88. Early Breast Cancer Trialists' Collaborative Group. Effects of radiotherapy and surgery in early breast cancer—an overview of the randomized trials. *N Engl J Med*. 1995;333:1444-1456.

89. Al-Ghazal S, Fallowfield L, Blamey R. Comparison of psychological aspects and patient satisfaction following breast conserving surgery, simple mastectomy and breast reconstruction. *Eur J Cancer*. 2000;36:1938-1943.

90. Schain WS, d'Angelo TM, Dunn ME, Lichter AS, Pierce LJ. Mastectomy versus conservative surgery and radiation therapy. Psychosocial consequences. *Cancer*. 1994;73:1221-1228.

91. Tang SS, Kaptanis S, Haddow JB, Mondani G, Elsberger B, Tasoulis MK*, et al.* Current margin practice and effect on re-excision rates following the publication of the SSO-ASTRO consensus and ABS consensus guidelines: a national prospective study of 2858 women undergoing breast-conserving therapy in the UK and Ireland. *Eur J Cancer*. 2017;84:315-324.

92. Houssami N, Macaskill P, Luke Marinovich M, Morrow M. The Association of Surgical Margins and Local Recurrence in Women with Early-Stage Invasive Breast Cancer Treated with Breast-Conserving Therapy: A Meta-Analysis. *Ann Surg Oncol*. 2014;21:717-730.

93. Fisher B, Jeong J-H, Anderson S, Bryant J, Fisher ER, Wolmark N. Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation. *N Engl J Med*. 2002;347:567-575.

94. Giuliano AE, McCall L, Beitsch P, Whitworth PW, Blumencranz P, Leitch AM*, et al.* Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. *Ann Surg*. 2010;252:426.

95. Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Costantino JP*, et al.* Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial. *Lancet Oncol*. 2010;11:927-933.

96. Krag D, Weaver D, Ashikaga T, Moffat F, Klimberg VS, Shriver C*, et al.* The sentinel node in breast cancer—a multicenter validation study. *N Engl J Med*. 1998;339:941-946.

97. The POSNOC trial. POsitive Sentinel NOde: adjuvant therapy alone versus adjuvant therapy plus Clearance or axillary radiotherapy [Internet]. NCTU POSNOC trial management team; 2019 [Accessed 04/04/2019]. Available from: http://www.posnoc.co.uk/.

98. Cutress RI, McIntosh SA, Potter S, Goyal A, Kirwan CC, Harvey J*, et al.* Opportunities and priorities for breast surgical research. *Lancet Oncol*. 2018;19:e521-e533.
99. Raman CV, Krishnan KS. A new type of secondary radiation. *Nature*. 1928;121:501.
100. Smith E, Dent G. Modern Raman Spectroscopy. 1st ed. Wiley, editor. Chichester: John Wiley & Sons Ltd; 2005.
101. Larkin PJ. Infrared and Raman spectroscopy : principles and spectral interpretation / , Spectroscopy and Materials Characterization. In: Salvary S, CT, United States, editor. 2018. p. 0-1.
102. Soderberg T. Organic chemistry with a biological emphasis Volume I. Chemistry Publications: Timothy Soderberg; 2016.
103. Movasaghi Z, Rehman S, Rehman IU. Raman Spectroscopy of Biological Tissues. *Appl Spectrosc Rev*. 2007;42:493-541.
104. Shipp DW, Sinjab F, Notingher I. Raman spectroscopy: techniques and applications in the life sciences. *Adv Opt Photonics*. 2017;9:315-428.
105. Mourant JR, Short KW, Carpenter S, Kunapareddy N, Coburn L, Powers T*, et al.* Biochemical differences in tumorigenic and nontumorigenic cells measured by Raman and infrared spectroscopy. *J Biomed Opt*. 2005;10:031106.
106. Kuhar N, Sil S, Verma T, Umapathy S. Challenges in application of Raman spectroscopy to biology and materials. *RSC advances*. 2018;8:25888-25908.
107. Aubertin K, Desroches J, Jermyn M, Trinh VQ, Saad F, Trudel D*, et al.* Combining high wavenumber and fingerprint Raman spectroscopy for the detection of prostate cancer during radical prostatectomy. *Biomed Opt Express*. 2018;9:4294 - 4305.
108. Manoharan R, Wang Y, Boustany NN, Brennan JF, Baraga JJ, Dasari RR*, et al.*, editors. Raman spectroscopy for cancer detection: instrument development and tissue diagnosis. Biomedical Optoelectronic Devices and Systems II; 1994: International Society for Optics and Photonics.
109. Notingher I, Verrier S, Romanska H, Bishop A, Polak J, Hench L. In situ characterisation of living cells by Raman spectroscopy. *J Spectrosc*. 2002;16:43-51.
110. Pascut FC, Goh HT, Welch N, Buttery LD, Denning C, Notingher I. Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells. *Biophys J*. 2011;100:251-259.

111. McCreery RL. Chapter 8: Dispersive Raman Spectrometers. In: Winefordner JD, editor. Raman Spectroscopy for Chemical Analysis. 1st ed. New York: John Wiley & Sons; 2000. p. 149-221.

112. Princeton Instruments. Introduction to scientific InGaAs FPA cameras [Internet]. Princeton Instruments; 2019 [Accessed 05/12/2019]. Available from: https://www.princetoninstruments.com/userfiles/files/technotes/Introduction-to-scientific-InGaAs-FPA-cameras.pdf.

113. Hubbard TJE, Shore A, Stone N. Raman spectroscopy for rapid intra-operative margin analysis of surgically excised tumour specimens. *Analyst*. 2019;144:6479-6496.

114. Gage I, Schnitt SJ, Nixon AJ, Silver B, Recht A, Troyan SL, *et al.* Pathologic margin involvement and the risk of recurrence in patients treated with breast-conserving therapy. *Cancer*. 1996;78:1921-1928.

115. Neuzillet Y, Soulie M, Larre S, Roupret M, Defortescu G, Murez T, *et al.* Positive surgical margins and their locations in specimens are adverse prognosis features after radical cystectomy in non-metastatic carcinoma invading bladder muscle: results from a nationwide case-control study. *BJU Int*. 2013;111:1253-1260.

116. Partridge M, Li S-R, Pateromichelakis S, Francis R, Phillips E, Huang XH, *et al.* Detection of minimal residual cancer to investigate why oral tumors recur despite seemingly adequate treatment. *Clin Cancer Res*. 2000;6:2718-2725.

117. Grant Y, Al-Khudairi R, St John E, Carter A, Barschkett M, Cunningham D, *et al.* Patient level costs of margin excision and re-excision for breast conserving surgery. *Eur J Surg Oncol*. 2018;44:891.

118. Yu J, Elmore LC, Cyr AE, Aft RL, Gillanders WE, Margenthaler JA. Cost analysis of a surgical consensus guideline in breast-conserving surgery. *J Am Coll Surg*. 2017;225:294-301.

119. Raab SS, Grzybicki DM, Janosky JE, Zarbo RJ, Meier FA, Jensen C, *et al.* Clinical impact and frequency of anatomic pathology errors in cancer diagnoses. *Cancer*. 2005;104:2205-2213.

120. Elmore JG, Nelson HD, Pepe MS, Longton GM, Tosteson AN, Geller B, *et al.* Variability in pathologists' interpretations of individual breast biopsy slides: a population perspective. *Ann Intern Med*. 2016;164:649-655.

121. Wazer DE, Schmidt-Ullrich RK, Schmid CH, Ruthazer R, Kramer B, Safaii H*, et al.* The value of breast lumpectomy margin assessment as a predictor of residual tumor burden. *Int J Radiat Oncol Biol Phys*. 1997;38:291-299.

122. Woolgar JA, Triantafyllou A. Pitfalls and procedures in the histopathological diagnosis of oral and oropharyngeal squamous cell carcinoma and a review of the role of pathology in prognosis. *Oral Oncol*. 2009;45:361-385.

123. Silberstein JL, Eastham JA. Significance and management of positive surgical margins at the time of radical prostatectomy. *Indian J Urol*. 2014;30:423-428.

124. Alicandri-Ciufelli M, Bonali M, Piccinini A, Marra L, Ghidini A, Cunsolo EM*, et al.* Surgical margins in head and neck squamous cell carcinoma: what is 'close'? *Eur Arch Otorhinolaryngol*. 2013;270:2603-2609.

125. Jorns JM, Daignault S, Sabel MS, Wu AJ. Is intraoperative frozen section analysis of reexcision specimens of value in preventing reoperation in breast-conserving therapy? *Am J Clin Pathol*. 2014;142:601-608.

126. Smeets NW, Krekels GA, Ostertag JU, Essers BA, Dirksen CD, Nieman FH*, et al.* Surgical excision vs Mohs' micrographic surgery for basal-cell carcinoma of the face: randomised controlled trial. *Lancet*. 2004;364:1766-1772.

127. Du E, Ow TJ, Lo YT, Gersten A, Schiff BA, Tassler AB*, et al.* Refining the utility and role of Frozen section in head and neck squamous cell carcinoma resection. *Laryngoscope*. 2016;126:1768-1775.

128. Thomas Robbins K, Triantafyllou A, Suarez C, Lopez F, Hunt JL, Strojan P*, et al.* Surgical margins in head and neck cancer: Intra- and postoperative considerations. *Auris Nasus Larynx*. 2018.

129. Weber WP, Engelberger S, Viehl CT, Zanetti-Dallenbach R, Kuster S, Dirnhofer S*, et al.* Accuracy of frozen section analysis versus specimen radiography during breast-conserving surgery for nonpalpable lesions. *World J Surg*. 2008;32:2599.

130. Olson T, Harter J, MuÒoz A, Mahvi D, Breslin TM. Frozen Section Analysis for Intraoperative Margin Assessment During Breast-Conserving Surgery Results in Low Rates of Re-excision and Local Recurrence. *Ann Surg Oncol*. 2007;14:2953-2960.

131. Tsuboi T, Ohori M, Kuroiwa K, Reuter VE, Kattan MW, Eastham JA*, et al.* Is intraoperative frozen section analysis an efficient way to reduce positive surgical margins? *Urology*. 2005;66:1287-1291.

132. Osborn JB, Keeney GL, Jakub JW, Degnim AC, Boughey JC. Cost-Effectiveness Analysis of Routine Frozen-Section Analysis of Breast Margins Compared with Reoperation for Positive Margins. *Ann Surg Oncol*. 2011;18:3204.

133. Nunez AL, Giannico GA, Mukhtar F, Dailey V, El-Galley R, Hameed O. Frozen section evaluation of margins in radical prostatectomy specimens: A contemporary study and literature review. *Ann Diagn Pathol*. 2016;24:11-18.

134. D'Halluin F, Tas P, Rouquette S, Bendavid C, Foucher F, Meshba H*, et al.* Intra-operative touch preparation cytology following lumpectomy for breast cancer: a series of 400 procedures. *Breast J*. 2009;18:248-253.

135. Simiyoshi K, Nohara T, Iwamoto M, Tanaka S, Kimura K, Takahashi Y*, et al.* Usefulness of intraoperative touch smear cytology in breast-conserving surgery. *Exp Ther Med*. 2010;1:641-645.

136. St John ER, Al-Khudairi R, Ashrafian H, Athanasiou T, Takats Z, Hadjiminas DJ*, et al.* Diagnostic Accuracy of Intraoperative Techniques for Margin Assessment in Breast Cancer Surgery. *Ann Surg*. 2017;265:300-310.

137. Mesurolle B, El-Khoury M, Hori D, Phancao J-P, Kary S, Kao E*, et al.* Sonography of Postexcision Specimens of Nonpalpable Breast Lesions: Value, Limitations, and Description of a Method. *AJR Am J Roentgenol*. 2006;186:1014-1024.

138. Londero V, Zuiani C, Panozzo M, Linda A, Girometti R, Bazzocchi M. Surgical specimen ultrasound: Is it able to predict the status of resection margins after breast-conserving surgery? *Breast J*. 2010;19:532-537.

139. Sastry R, Bi WL, Pieper S, Frisken S, Kapur T, Wells III W*, et al.* Applications of ultrasound in the resection of brain tumors. *J Neuroimaging*. 2017;27:5-15.

140. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. *Lancet Oncol*. 2011;12:997-1003.

141. Sanai N, Berger MS. Surgical oncology for gliomas: the state of the art. *Nat Rev Clin Oncol*. 2018;15:112 - 125.

142. Barone DG, Lawrie TA, Hart MG. Image guided surgery for the resection of brain tumours. *Cochrane Database Syst Rev*. 2014.

143. Glaser AK, Reder NP, Chen Y, McCarty EF, Yin C, Wei L*, et al.* Light-sheet microscopy for slide-free non-destructive pathology of large clinical specimens. *Nat Biomed Eng*. 2017;1:0084.

144. Fereidouni F, Harmany ZT, Tian M, Todd A, Kintner JA, McPherson JD*, et al.* Microscopy with ultraviolet surface excitation for rapid slide-free histology. *Nat Biomed Eng*. 2017;1:957.

145. Nguyen FT, Zysk AM, Chaney EJ, Kotynek JG, Oliphant UJ, Bellafiore FJ*, et al.* Intraoperative evaluation of breast tumor margins with optical coherence tomography. *Cancer Res*. 2009;69:8790-8796.

146. Zysk AM, Chen K, Gabrielson E, Tafra L, Gonzalez EAM, Canner JK*, et al.* Intraoperative assessment of final margins with a handheld optical imaging probe during breast-conserving surgery may reduce the reoperation rate: results of a multicenter study. *Ann Surg Oncol*. 2015;22:3356-3362.

147. Nachabé R, Hendriks BH, Lucassen GW, van der Voort M, Evers DJ, Rutgers EJ*, et al.* Diagnosis of breast cancer using diffuse optical spectroscopy from 500 to 1600 nm: comparison of classification methods. *J Biomed Opt*. 2011;16:087010.

148. Wilke LG, Brown JQ, Bydlon TM, Kennedy SA, Richards LM, Junker MK*, et al.* Rapid noninvasive optical imaging of tissue composition in breast tumor margins. *Am J Surg*. 2009;198:566-574.

149. Chase B. A new generation of Raman instrumentation. *J Appl Spectrosc*. 1994;48:14A-19A.

150. Delhaye M, Dhamelincourt P. Raman microprobe and microscope with laser excitation. *J Raman Spectrosc*. 1975;3:33-43.

151. De Grauw CJ, Otto C, Greve J. Line-Scan Raman Microspectrometry for Biological Applications. *Appl Spectrosc*. 1997;51:1607-1612.

152. Hutchings J, Kendall C, Smith B, Shepherd N, Barr H, Stone N. The potential for histological screening using a combination of rapid Raman mapping and principal component analysis. *J Biophotonics*. 2009;2:91-103.

153. Sinjab F, Kong K, Gibson G, Varma S, Williams H, Padgett M*, et al.* Tissue diagnosis using power-sharing multifocal Raman micro-spectroscopy and auto-fluorescence imaging. *Biomed Opt Express*. 2016;7:2993-3006.

154. Schlücker S, Schaeberle MD, Huffman SW, Levin IW. Raman Microspectroscopy: A Comparison of Point, Line, and Wide-Field Imaging Methodologies. *Anal Chem*. 2003;75:4312-4318.

155. Rowlands CJ, Varma S, Perkins W, Leach I, Williams H, Notingher I. Rapid acquisition of Raman spectral maps through minimal sampling: applications in tissue imaging. *J Biophotonics*. 2012;5:220-229.

156. Kong K, Zaabar F, Rakha E, Ellis I, Koloydenko A, Notingher I. Towards intra-operative diagnosis of tumours during breast conserving surgery by selective-sampling Raman micro-spectroscopy. *Phys Med Biol*. 2014;59:6141-6152.

157. Leigh SY, Som M, Liu JTC. Method for Assessing the Reliability of Molecular Diagnostics Based on Multiplexed SERS-Coded Nanoparticles. *PLoS ONE*. 2013;8:e62084.

158. Laing S, Jamieson LE, Faulds K, Graham D. Surface-enhanced Raman spectroscopy for in vivo biosensing. *Nat Rev Chem*. 2017;1:0060.

159. Yezhelyev MV, Al-Hajj A, Morris C, Marcus AI, Liu T, Lewis M*, et al.* In situ molecular profiling of breast cancer biomarkers with multicolor quantum dots. *Adv Mater*. 2007;19:3146-3151.

160. Matousek P, Clark IP, Draper ERC, Morris MD, Goodship AE, Everall N*, et al.* Subsurface Probing in Diffusely Scattering Media Using Spatially Offset Raman Spectroscopy. *Appl Spectrosc*. 2005;59:393-400.

161. Schulmerich MV, Dooley K, Morris MD, Vanasse TM, Goldstein SA. Transcutaneous fiber optic Raman spectroscopy of bone using annular illumination and a circular array of collection fibers. *J Biomed Opt*. 2006;11:060502.

162. Ghita A, Matousek P, Stone N. High sensitivity non-invasive detection of calcifications deep inside biological tissue using Transmission Raman Spectroscopy. *J Biophotonics*. 2018;11:e201600260.

163. Santos LF, Wolthuis R, Koljenović S, Almeida RM, Puppels GJ. Fiber-Optic Probes for in Vivo Raman Spectroscopy in the High-Wavenumber Region. *Anal Chem*. 2005;77:6747-6752.

164. Petterson IEI, Day JC, Fullwood LM, Gardner B, Stone N. Characterisation of a fibre optic Raman probe within a hypodermic needle. *Anal Bioanal Chem*. 2015;407:8311-8320.

165. Desroches J, Jermyn M, Pinto M, Picot F, Tremblay MA, Obaid S*, et al.* A new method using Raman spectroscopy for in vivo targeted brain cancer tissue biopsy. *Sci Rep*. 2018;8:1792.

166. Day J, Bennett R, Smith B, Kendall C, Hutchings J, Meaden G*, et al.* A miniature confocal Raman probe for endoscopic use. *Phys Med Biol*. 2009;54:7077.

167. Pahlow S, Weber K, Popp J, Wood BR, Kochan K, Ruther A*, et al.* Application of Vibrational Spectroscopy and Imaging to Point-of-Care Medicine: A Review. *Appl Spectrosc*. 2018;72:52-84.

168. Day JCC, Stone N. A Subcutaneous Raman Needle Probe. *Appl Spectrosc*. 2013;67:349-354.

169. Stevens O, Iping Petterson IE, Day JC, Stone N. Developing fibre optic Raman probes for applications in clinical spectroscopy. *Chem Soc Rev*. 2016;45:1919-1934.

170. Camp Jr CH, Cicerone MT. Chemically sensitive bioimaging with coherent Raman scattering. *Nat Photonics*. 2015;9:295-305.

171. Evans CL, Potma EO, Puoris' haag M, Côté D, Lin CP, Xie XS. Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy. *Proc Natl Acad Sci U S A*. 2005;102:16807-16812.

172. Saar BG, Freudiger CW, Reichman J, Stanley CM, Holtom GR, Xie XS. Video-rate molecular imaging in vivo with stimulated Raman scattering. *Science*. 2010;330:1368-1370.

173. Ji M, Lewis S, Camelo-Piragua S, Ramkissoon SH, Snuderl M, Venneti S*, et al.* Detection of human brain tumor infiltration with quantitative stimulated Raman scattering microscopy. *Sci Transl Med*. 2015;7:309ra163-309ra163.

174. Barroso E, Smits R, Bakker Schut T, Ten Hove I, Hardillo J, Wolvius E*, et al.* Discrimination between oral cancer and healthy tissue based on water content determined by Raman spectroscopy. *Anal Chem*. 2015;87:2419-2426.

175. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci*. 2003;43:1947-1958.

176. Dingari NC, Barman I, Myakalwar AK, Tewari SP, Kumar Gundawar M. Incorporation of support vector machines in the LIBS toolbox for sensitive and robust classification amidst unexpected sample and system variability. *Anal Chem*. 2012;84:2686-2694.

177. Ciccarelli G, Di Virgilio M, Menna S, Garretti L, Ala A, Giani R*, et al.* Radiography of the surgical specimen in early stage breast lesions: diagnostic reliability in the analysis of the resection margins. *Radiol Med*. 2007;112:366-376.

178. Moran MS, Schnitt SJ, Giuliano AE, Harris JR, Khan SA, Horton J*, et al.* Society of Surgical Oncology–American Society for Radiation Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in stages I and II invasive breast cancer. *Ann Surg Oncol*. 2014;21:704-716.

179. Deng K, Zhu C, Ma X, Jia H, Wei Z, Xiao Y*, et al.* Rapid Discrimination of Malignant Breast Lesions from Normal Tissues Utilizing Raman Spectroscopy System: A Systematic Review and Meta-Analysis of In Vitro Studies. *PLoS One*. 2016;11:e0159860.

180. Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, Feld MS. Diagnosing breast cancer by using Raman spectroscopy. *Proc Natl Acad Sci U S A*. 2005;102:12371-12376.

181. Majumder SK, Keller MD, Boulos FI, Kelley MC, Mahadevan-Jansen A. Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination. *J Biomed Opt*. 2008;13:054009.

182. Barman I, Dingari NC, Saha A, McGee S, Galindo LH, Liu W*, et al.* Application of Raman spectroscopy to identify microcalcifications and underlying breast lesions at stereotactic core needle biopsy. *Cancer Res*. 2013;73:3206-3215.

183. Baker R, Rogers K, Shepherd N, Stone N. New relationships between breast microcalcifications and cancer. *Br J Cancer*. 2010;103:1034.

184. Han B, Du Y, Fu T, Fan Z, Xu S, Hu C*, et al.* Differences and Relationships Between Normal and Atypical Ductal Hyperplasia, Ductal Carcinoma In Situ, and Invasive Ductal Carcinoma Tissues in the Breast Based on Raman Spectroscopy. *Appl Spectrosc*. 2017;71:300-307.

185. Keller MD, Vargis E, de Matos Granja N, Wilson RH, Mycek MA, Kelley MC*, et al.* Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation. *J Biomed Opt*. 2011;16:077006.

186. Haka AS, Volynskaya Z, Gardecki JA, Nazemi J, Shenk R, Wang N*, et al.* Diagnosing breast cancer using Raman spectroscopy: prospective analysis. *J Biomed Opt*. 2009;14:054023.

187. Langhans L, Jensen M, Talman MM, Vejborg I, Kroman N, Tvedskov TF. Reoperation rates in ductal carcinoma in situ vs invasive breast cancer after wire-guided breast-conserving surgery. *JAMA Surg*. 2017;152:378-384.

188. Wang YW, Reder NP, Kang S, Glaser AK, Yang Q, Wall MA*, et al.* Raman-Encoded Molecular Imaging with Topically Applied SERS Nanoparticles for Intraoperative Guidance of Lumpectomy. *Cancer Res*. 2017;77:4506-4516.

189. Shipp DW, Rakha EA, Koloydenko AA, Macmillan RD, Ellis IO, Notingher I. Intra-operative spectroscopic assessment of surgical margins during breast conserving surgery. *Breast Cancer Res*. 2018;20:69.

190. Guy Jr GP, Machlin SR, Ekwueme DU, Yabroff KR. Prevalence and Costs of Skin Cancer Treatment in the US, 2002− 2006 and 2007− 2011. *Am J Prev Med*. 2015;48:183-187.

191. Gulleth Y, Goldberg N, Silverman RP, Gastman BR. What is the best surgical margin for a basal cell carcinoma: a meta-analysis of the literature. *Plast Reconstr Surg*. 2010;126:1222-1231.

192. Telfer N, Colver G, Morton C. Guidelines for the management of basal cell carcinoma. *Br J Dermatol*. 2008;159:35-48.

193. Mosterd K, Krekels GAM, Nieman FHM, Ostertag JU, Essers BAB, Dirksen CD*, et al.* Surgical excision versus Mohs' micrographic surgery for primary and recurrent basal-cell carcinoma of the face: a prospective randomised controlled trial with 5-years' follow-up. *Lancet Oncol*. 2008;9:1149-1156.

194. Nikolaou V, Stratigos A. Emerging trends in the epidemiology of melanoma. *Br J Dermatol*. 2014;170:11-19.

195. National Institute of Health and Care Excellence (NICE). Melanoma: assessment and management [NG14] [Internet]. NICE; 2015 [Accessed 29/01/2019]. Available from: https://www.nice.org.uk/guidance/ng14.

196. Sladden MJ, Balch C, Barzilai DA, Berg D, Freiman A, Handiside T*, et al.* Surgical excision margins for primary cutaneous melanoma. *Cochrane Database Syst Rev*. 2009;4:CD004835.

197. Schut TCB, Caspers PJ, Puppels GJ, Nijssen A, Heule F, Neumann MH*, et al.* Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy. *J Invest Dermatol*. 2002;119:64-69.

198. Nijssen A, Maquelin K, Santos LF, Caspers PJ, Schut TCB, den Hollander JC*, et al.* Discriminating basal cell carcinoma from perilesional skin using high wave-number Raman spectroscopy. *J Biomed Opt*. 2007;12:034004.

199. Lieber CA, Majumder SK, Ellis DL, Billheimer DD, Mahadevan-Jansen A. In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy. *Lasers Surg Med*. 2008;40:461-467.

200. Lieber CA, Majumder SK, Billheimer DD, Ellis DL, Mahadevan-Jansen A. Raman microspectroscopy for skin cancer detection in vitro. *J Biomed Opt*. 2009;13:024013.

201. Schleusener J, Gluszczynska P, Reble C, Gersonde I, Helfmann J, Fluhr JW*, et al.* In vivo study for the discrimination of cancerous and normal skin using fibre probe-based Raman spectroscopy. *Exp Dermatol*. 2015;24:767-772.

202. Lui H, Zhao J, McLean DI, Zeng H. Real-time Raman spectroscopy for in vivo skin cancer diagnosis. *Cancer Res*. 2012;72:2491-2500.

203. Zhao J, Lui H, Kalia S, Zeng H. Real-time Raman spectroscopy for automatic in vivo skin cancer detection: an independent validation. *Anal Bioanal Chem*. 2015;407:8373-8379.

204. Zhang J, Fan Y, Song Y, Xu J. Accuracy of Raman spectroscopy for differentiating skin cancer from normal tissue. *Medicine (Baltimore)*. 2018;97:e12022.

205. Bratchenko IA, Artemyev DN, Myakinin OO, Khristoforova YA, Moryatov AA, Kozlov SV*, et al.* Combined Raman and autofluorescence ex vivo diagnostics of skin cancer in near-infrared and visible regions. *J Biomed Opt*. 2017;22:027005.

206. Heuke S, Vogler N, Meyer T, Akimov D, Kluschke F, Röwert-Huber HJ*, et al.* Multimodal mapping of human skin. *Br J Dermatol*. 2013;169:794-803.

207. Kong K, Rowlands CJ, Varma S, Perkins W, Leach IH, Koloydenko AA*, et al.* Diagnosis of tumors during tissue-conserving surgery with integrated autofluorescence and Raman scattering microscopy. *Proc Natl Acad Sci U S A*. 2013;110:15189-15194.

208. Kong K, Rowlands CJ, Varma S, Perkins W, Leach IH, Koloydenko AA*, et al.* Increasing the speed of tumour diagnosis during surgery with selective scanning Raman microscopy. *J Mol Struct*. 2014;1073:58-65.

209. Boitor R, Kong K, Shipp D, Varma S, Koloydenko A, Kulkarni K*, et al.* Automated multimodal spectral histopathology for quantitative diagnosis of residual tumour during basal cell carcinoma surgery. *Biomed Opt Express*. 2017;8:5749-5766.

210. Goodenberger ML, Jenkins RB. Genetics of adult glioma. *Cancer Genet*. 2012;205:613-621.

211. National Institute for Health and Care Excellence (NICE). Brain Tumours (primary) and brain metastases in adults [NG99] [Internet]. NICE; 2018

[Accessed 24/01/2019]. Available from: www.nice.org.uk/guidance/ng99/chapter/recommendations#management-of-glioma.

212. Petrecca K, Guiot M-C, Panet-Raymond V, Souhami L. Failure pattern following complete resection plus radiotherapy and temozolomide is at the resection margin in patients with glioblastoma. *J Neurooncol*. 2013;111:19-23.

213. Beljebbar A, Dukic S, Amharref N, Manfait M. Ex vivo and in vivo diagnosis of C6 glioblastoma development by Raman spectroscopy coupled to a microprobe. *Anal Bioanal Chem*. 2010;398:477-487.

214. Ji M, Orringer DA, Freudiger CW, Ramkissoon S, Liu X, Lau D*, et al.* Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy. *Sci Transl Med*. 2013;5:201ra119-201ra119.

215. Uckermann O, Galli R, Tamosaityte S, Leipnitz E, Geiger KD, Schackert G*, et al.* Label-free delineation of brain tumors by coherent anti-stokes Raman scattering microscopy in an orthotopic mouse model and human glioblastoma. *PLoS One*. 2014;9:e107115.

216. Gao X, Yue Q, Liu Z, Ke M, Zhou X, Li S*, et al.* Guiding Brain-Tumor Surgery via Blood–Brain-Barrier-Permeable Gold Nanoprobes with Acid-Triggered MRI/SERRS Signals. *Adv Mater*. 2017;29:1603917.

217. Karabeber H, Huang R, Iacono P, Samii JM, Pitter K, Holland EC*, et al.* Guiding brain tumor resection using surface-enhanced Raman scattering nanoparticles and a hand-held Raman scanner. *ACS Nano*. 2014;8:9755-9766.

218. Demeule M, Currie JC, Bertrand Y, Ché C, Nguyen T, Régina A*, et al.* Involvement of the low-density lipoprotein receptor-related protein in the transcytosis of the brain delivery vector Angiopep-2. *J Neurochem*. 2008;106:1534-1544.

219. Bergner N, Bocklitz T, Romeike BFM, Reichart R, Kalff R, Krafft C*, et al.* Identification of primary tumors of brain metastases by Raman imaging and support vector machines. *Chemometr Intell Lab Syst*. 2012;117:224-232.

220. Kalkanis SN, Kast RE, Rosenblum ML, Mikkelsen T, Yurgelevic SM, Nelson KM*, et al.* Raman spectroscopy to distinguish grey matter, necrosis, and glioblastoma multiforme in frozen tissue sections. *J Neurooncol*. 2014;116:477-485.

221. Orringer DA, Pandian B, Niknafs YS, Hollon TC, Boyle J, Lewis S*, et al.* Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nat Biomed Eng*. 2017;1:0027.

222. Bury D, Morais CL, Ashton KM, Dawson TP, Martin FL. Ex Vivo Raman Spectrochemical Analysis Using a Handheld Probe Demonstrates High Predictive Capability of Brain Tumour Status. *Biosensors*. 2019;9:49.

223. Jermyn M, Mok K, Mercier J, Desroches J, Pichette J, Saint-Arnaud K*, et al.* Intraoperative brain cancer detection with Raman spectroscopy in humans. *Sci Transl Med*. 2015;7:274ra219.

224. Jermyn M, Mercier J, Aubertin K, Desroches J, Urmey K, Karamchandiani J*, et al.* Highly Accurate Detection of Cancer In Situ with Intraoperative, Label-Free, Multimodal Optical Spectroscopy. *Cancer Res*. 2017;77:3942-3950.

225. National Institute for Health Care and Excellence (NICE). Cancer of the upper aerodigestive tract: assessment and management in people aged 16 and over [NG36] [Internet]. NICE; 2016 [Accessed 24/01/2019]. Available from: https://www.nice.org.uk/guidance/ng36.

226. Loree TR, Strong EW. Significance of positive margins in oral cavity squamous carcinoma. *Am J Surg*. 1990;160:410-414.

227. Smits RW, Koljenović S, Hardillo JA, ten Hove I, Meeuwis CA, Sewnaik A*, et al.* Resection margins in oral cancer surgery: room for improvement. *Head Neck*. 2016;38:E2197-E2203.

228. Barroso EM, Smits RW, van Lanschot CG, Caspers PJ, Ten Hove I, Mast H*, et al.* Water Concentration Analysis by Raman Spectroscopy to Determine the Location of the Tumor Border in Oral Cancer Surgery. *Cancer Res*. 2016;76:5945-5953.

229. Helliwell J, Woolgar T. Standards and datasets for reporting cancers Dataset for histopathology reporting of mucosal malignancies of the oral cavity [Internet]. Royal College of Pathologists; 2013 [Accessed 24/01/2019]. Available from: https://www.rcpath.org/uploads/assets/uploaded/e485236e-6d2a-4109-96fe1da9d5e140f9.pdf.

230. Meier JD, Oliver DA, Varvares MA. Surgical margin determination in head and neck oncology: current clinical practice. The results of an International American Head and Neck Society Member Survey. *Head Neck*. 2005;27:952-958.

231. Pathak KA, Nason RW, Penner C, Viallet NR, Sutherland D, Kerr PD. Impact of use of frozen section assessment of operative margins on survival in oral cancer. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2009;107:235-239.

232. Stone N, Stavroulaki P, Kendall C, Birchall M, Barr H. Raman spectroscopy for early detection of laryngeal malignancy: preliminary results. *Laryngoscope*. 2000;110:1756-1763.

233. Lau DP, Huang Z, Lui H, Anderson DW, Berean K, Morrison MD*, et al.* Raman spectroscopy for optical diagnosis in the larynx: preliminary findings. *Lasers Surg Med*. 2005;37:192-200.

234. Lin K, Cheng DLP, Huang Z. Optical diagnosis of laryngeal cancer using high wavenumber Raman spectroscopy. *Biosens Bioelectron*. 2012;35:213-217.

235. Lin K, Zheng W, Lim CM, Huang Z. Real-time in vivo diagnosis of laryngeal carcinoma with rapid fiber-optic Raman spectroscopy. *Biomed Opt Express*. 2016;7:3705-3715.

236. Lin K, Zheng W, Lim CM, Huang Z. Real-time in vivo diagnosis of nasopharyngeal carcinoma using rapid fiber-optic Raman spectroscopy. *Theranostics*. 2017;7:3517-3526.

237. Cals FLJ, Bakker Schut TC, Hardillo JA, Baatenburg de Jong RJ, Koljenović S, Puppels GJ. Investigation of the potential of Raman spectroscopy for oral cancer detection in surgical margins. *Lab Invest*. 2015;95:1186-1196.

238. Cals FLJ, Koljenović S, Hardillo JA, Baatenburg de Jong RJ, Bakker Schut TC, Puppels GJ. Development and validation of Raman spectroscopic classification models to discriminate tongue squamous cell carcinoma from non-tumorous tissue. *Oral Oncol*. 2016;60:41-47.

239. Guze K, Pawluk HC, Short M, Zeng H, Lorch J, Norris C*, et al.* Pilot study: Raman spectroscopy in differentiating premalignant and malignant oral lesions from normal mucosa and benign lesions in humans. *Head Neck*. 2015;37:511-517.

240. Barroso EM, ten Hove I, Bakker Schut TC, Mast H, van Lanschot CGF, Smits RWH*, et al.* Raman spectroscopy for assessment of bone resection margins in mandibulectomy for oral cavity squamous cell carcinoma. *Eur J Cancer*. 2018;92:77-87.

241. Maheedhar K, Bhat RA, Malini R, Prathima N, Keerthi P, Kushtagi P*, et al.* Diagnosis of ovarian cancer by Raman spectroscopy: a pilot study. *Photomed Laser Surg*. 2008;26:83-90.

242. McGregor HC, Short MA, McWilliams A, Shaipanich T, Ionescu DN, Zhao J*, et al.* Real-time endoscopic Raman spectroscopy for in vivo early lung cancer detection. *J Biophotonics*. 2017;10:98-110.

243. Gao L, Wang Z, Li F, Hammoudi AA, Thrall MJ, Cagle PT*, et al.* Differential diagnosis of lung carcinoma with coherent anti-Stokes Raman scattering imaging. *Arch Pathol Lab Med*. 2012;136:1502-1510.

244. Teixeira CS, Bitar RA, Martinho HS, Santos AB, Kulcsar MA, Friguglietti CU*, et al.* Thyroid tissue analysis through Raman spectroscopy. *Analyst*. 2009;134:2361-2370.

245. Crow P, Stone N, Kendall CA, Uff JS, Farmer JAM, Barr H*, et al.* The use of Raman spectroscopy to identify and grade prostatic adenocarcinoma in vitro. *Br J Cancer*. 2003;89:106-108.

246. Stone N, Hart Prieto MC, Crow P, Uff J, Ritchie AW. The use of Raman spectroscopy to provide an estimation of the gross biochemistry associated with urological pathologies. *Anal Bioanal Chem*. 2007;387:1657-1668.

247. Aubertin K, Birlea M, Vladoiu MC, Grosset AA, Trinh VQ, Trudel D*, et al.* Mesoscopic characterization of prostate cancer using Raman spectroscopy: potential for diagnostics and therapeutics. *BJU Int*. 2018;122:326-336.

248. Pinto M, Zorn KC, Tremblay J-P, Desroches J, Dallaire F, Aubertin K*, et al.* Integration of a Raman spectroscopy system to a robotic-assisted surgical system for real-time tissue characterization during radical prostatectomy procedures. *J Biomed Opt*. 2019;24:025001.

249. Robichaux-Viehoever A, Kanter E, Shappell H, Billheimer D, Jones H, 3rd, Mahadevan-Jansen A. Characterization of Raman spectra measured in vivo for the detection of cervical dysplasia. *Appl Spectrosc*. 2007;61:986-993.

250. Shaikh RS, Dora TK, Chopra S, Maheshwari A, Kedar DK, Bharat R*, et al.* In vivo Raman spectroscopy of human uterine cervix: exploring the utility of vagina as an internal control. *J Biomed Opt*. 2014;19:087001.

251. Duraipandian S, Zheng W, Ng J, Low JJ, Ilancheran A, Huang Z. Near-infrared-excited confocal Raman spectroscopy advances in vivo diagnosis of cervical precancer. *J Biomed Opt*. 2013;18:067007.

252. Crow P, Molckovsky A, Stone N, Uff J, Wilson B, WongKeeSong LM. Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer. *Urology*. 2005;65:1126-1130.

253. Draga RO, Grimbergen MC, Vijverberg PL, van Swol CF, Jonges TG, Kummer JA*, et al.* In vivo bladder cancer diagnosis by high-volume Raman spectroscopy. *Anal Chem*. 2010;82:5993-5999.

254. Kendall C, Day J, Hutchings J, Smith B, Shepherd N, Barr H*, et al.* Evaluation of Raman probe for oesophageal cancer diagnostics. *Analyst*. 2010;135:3038-3041.

255. Teh S, Zheng W, Ho K, Teh M, Yeoh K, Huang Z. Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach. *Br J Surg*. 2010;97:550-557.

256. Wang J, Lin K, Zheng W, Ho KY, Teh M, Yeoh KG*, et al.* Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy improves in vivo diagnosis of esophageal squamous cell carcinoma at endoscopy. *Sci Rep*. 2015;5:12957.

257. Bergholt MS, Lin K, Wang J, Zheng W, Xu H, Huang Q*, et al.* Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy enhances real-time in vivo diagnosis of adenomatous polyps during colonoscopy. *J Biophotonics*. 2016;9:333-342.

258. Gravante G, Hemingway D, Stephenson JA, Sharpe D, Osman A, Haines M*, et al.* Rectal cancers with microscopic circumferential resection margin involvement (R1 resections): survivals, patterns of recurrence, and prognostic factors. *J Surg Oncol*. 2016;114:642-648.

259. Dexter S, Sue-Ling H, McMahon M, Quirke P, Mapstone N, Martin I. Circumferential resection margin involvement: an independent predictor of survival following surgery for oesophageal cancer. *Gut*. 2001;48:667-670.

260. Quinn LM, Hollis AC, Hodson J, Elshafie MA, Hallissey MT, Whiting JL*, et al.* Prognostic significance of circumferential resection margin involvement in patients receiving potentially curative treatment for oesophageal cancer. *Eur J Surg Oncol*. 2018.

261. Pence I, Mahadevan-Jansen A. Clinical instrumentation and applications of Raman spectroscopy. *Chem Soc Rev*. 2016;45:1958-1979.

262. Haka AS, Volynskaya Z, Gardecki JA, Nazemi J, Lyons J, Hicks D*, et al.* In vivo margin assessment during partial mastectomy breast surgery using raman spectroscopy. *Cancer Res*. 2006;66:3317-3322.

263. Horsnell JD, Kendall C, Stone N. Towards the intra-operative use of Raman spectroscopy in breast cancer—overcoming the effects of theatre lighting. *Lasers Med Sci*. 2016;31:1143-1149.

264. Santos IP, Barroso EM, Bakker Schut TC, Caspers PJ, van Lanschot CGF, Choi DH*, et al.* Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics. *Analyst*. 2017;142:3025-3047.

265. Thomas G, Nguyen TQ, Pence IJ, Caldwell B, O'Connor ME, Giltnane J*, et al.* Evaluating feasibility of an automated 3-dimensional scanner using Raman spectroscopy for intraoperative breast margin assessment. *Sci Rep*. 2017;7:13548.

266. Dixon JM, Renshaw L, Young O, Kulkarni D, Saleem T, Sarfaty M*, et al.* Intra-operative assessment of excised breast tumour margins using ClearEdge imaging device. *Eur J Surg Oncol*. 2016;42:1834-1840.

267. Stables R, Clemens G, Butler HJ, Ashton KM, Brodbelt A, Dawson TP*, et al.* Feature driven classification of Raman spectra for real-time spectral brain tumour diagnosis using sound. *Analyst*. 2017;142:98-109.

268. Fink C, Haenssle H. Non-invasive tools for the diagnosis of cutaneous melanoma. *Skin Res Technol*. 2017;23:261-271.

269. Horsnell J, Stonelake P, Christie-Brown J, Shetty G, Hutchings J, Kendall C*, et al.* Raman spectroscopy—a new method for the intra-operative assessment of axillary lymph nodes. *Analyst*. 2010;135:3042-3047.

270. Wang Y, Kang S, Khan A, Ruttner G, Leigh SY, Murray M*, et al.* Quantitative molecular phenotyping with topically applied SERS nanoparticles for intraoperative guidance of breast cancer lumpectomy. *Sci Rep*. 2016;6:21242.

271. García-Flores A, Raniero L, Canevari R, Jalkanen KJ, Bitar R, Martinho H*, et al.* High-wavenumber FT-Raman spectroscopy for in vivo and ex vivo measurements of breast cancer. *Theor Chem Acc*. 2011;130:1231-1238.

272. Mo J, Zheng W, Low JJH, Ng J, Ilancheran A, Huang Z. High Wavenumber Raman Spectroscopy for in Vivo Detection of Cervical Dysplasia. *Anal Chem*. 2009;81:8908-8915.

273. Koljenovic S, Schut TCB, Wolthuis R, de Jong B, Santos L, Caspers PJ*, et al.* Tissue characterization using high wave number Raman spectroscopy. *J Biomed Opt*. 2005;10:031116.

274. Damadian R. Tumor detection by nuclear magnetic resonance. *Science*. 1971;171:1151-1153.

275. Sinha S, Lucas-Quesada FA, Sinha U, DeBruhl N, Bassett LW. In vivo diffusion-weighted MRI of the breast: potential for lesion characterization. *J Magn Reson Imaging*. 2002;15:693-704.

276. Cerussi A, Shah N, Hsiang D, Durkin A, Butler J, Tromberg BJ. In vivo absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy. *J Biomed Opt*. 2006;11:044005.

277. Verkman A, Hara-Chikuma M, Papadopoulos MC. Aquaporins—new players in cancer biology. *J Mol Med*. 2008;86:523-529.

278. Chung SH, Cerussi AE, Klifa C, Baek HM, Birgul O, Gulsen G*, et al.* In vivo water state measurements in breast cancer using broadband diffuse optical spectroscopy. *Phys Med Biol*. 2008;53:6713-6727.

279. Minard KR, Holtom GR, Kathmann LE, Majors PD, Thrall BD, Wind RA. Simultaneous 1H PFG-NMR and confocal microscopy of monolayer cell cultures: effects of apoptosis and necrosis on water diffusion and compartmentalization. *Magn Reson Med*. 2004;52:495-505.

280. Boyd N, Martin L, Chavez S, Gunasekara A, Salleh A, Melnichouk O*, et al.* Breast-tissue composition and other risk factors for breast cancer in young women: a cross-sectional study. *Lancet Oncol*. 2009;10:569-580.

281. Srinivasan S, Pogue BW, Jiang S, Dehghani H, Kogel C, Soho S*, et al.* Interpreting hemoglobin and water concentration, oxygen saturation, and scattering measured in vivo by near-infrared breast tomography. *Proc Natl Acad Sci U S A*. 2003;100:12349-12354.

282. Shah N, Cerussi AE, Jakubowski D, Hsiang D, Butler J, Tromberg BJ. Spatial variations in optical and physiological properties of healthy breast tissue. *J Biomed Opt*. 2004;9:534-540.

283. Taroni P, Paganoni AM, Ieva F, Pifferi A, Quarto G, Abbate F*, et al.* Non-invasive optical estimate of tissue composition to differentiate malignant from benign breast lesions: A pilot study. *Sci Rep*. 2017;7:40683.

284. Taroni P, Quarto G, Pifferi A, Abbate F, Balestreri N, Menna S*, et al.* Breast tissue composition and its dependence on demographic risk factors for breast cancer: non-invasive assessment by time domain diffuse optical spectroscopy. *PLoS One*. 2015;10:e0128941.

285. Graham S, Stanchev P, Lloyd-Smith J, Bronskill M, Plewes D. Changes in fibroglandular volume and water content of breast tissue during the menstrual cycle observed by MR imaging at 1.5 T. *J Magn Reson Imaging*. 1995;5:695-701.

286. Pogue BW, Jiang S, Dehghani H, Kogel C, Soho S, Srinivasan S*, et al.* Characterization of hemoglobin, water, and NIR scattering in breast tissue: analysis of intersubject variability and menstrual cycle changes. *J Biomed Opt*. 2004;9:541-552.

287. Fitzgerald AJ, Wallace VP, Jimenez-Linan M, Bobrow L, Pye RJ, Purushotham AD*, et al.* Terahertz pulsed imaging of human breast tumors. *Radiology*. 2006;239:533-540.

288. Grootendorst MR, Fitzgerald AJ, Brouwer de Koning SG, Santaolalla A, Portieri A, Van Hemelrijck M*, et al.* Use of a handheld terahertz pulsed imaging device to differentiate benign and malignant breast tissue. *Biomed Opt Express*. 2017;8:2932-2945.

289. Cerussi A, Hsiang D, Shah N, Mehta R, Durkin A, Butler J*, et al.* Predicting response to breast cancer neoadjuvant chemotherapy using diffuse optical spectroscopy. *Proc Natl Acad Sci U S A*. 2007;104:4014-4019.

290. Jagannathan N, Singh M, Govindaraju V, Raghunathan P, Coshic O, Julka P*, et al.* Volume localized in vivo proton MR spectroscopy of breast carcinoma: variation of water–fat ratio in patients receiving chemotherapy. *NMR Biomed*. 1998;11:414-422.

291. Agarwal K, Sharma U, Mathur S, Seenu V, Parshad R, Jagannathan NR. Study of lipid metabolism by estimating the fat fraction in different breast tissues and in various breast tumor sub-types by in vivo (1)H MR spectroscopy. *Magn Reson Imaging*. 2018;49:116-122.

292. de Boer LL, Hendriks BH, van Duijnhoven F, Peeters-Baas MT, Van de Vijver K, Loo CE*, et al.* Using DRS during breast conserving surgery: identifying robust optical parameters and influence of inter-patient variation. *Biomed Opt Express*. 2016;7:5188-5200.

293. de Boer LL, Molenkamp BG, Bydlon TM, Hendriks BH, Wesseling J, Sterenborg HJ*, et al.* Fat/water ratios measured with diffuse reflectance

spectroscopy to detect breast tumor boundaries. *Breast Cancer Res Treat*. 2015;152:509-518.

294. Wolthuis R, van Aken M, Fountas K, Robinson Jr JS, Bruining HA, Puppels GJ. Determination of water concentration in brain tissue by Raman spectroscopy. *Anal Chem*. 2001;73:3915-3920.

295. Caspers PJ, Lucassen GW, Carter EA, Bruining HA, Puppels GJ. In vivo confocal Raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles. *J Invest Dermatol*. 2001;116:434-442.

296. Huizinga A, Bot AC, de Mul FF, Vrensen GF, Greve J. Local variation in absolute water content of human and rabbit eye lenses measured by Raman microspectroscopy. *Exp Eye Res*. 1989;48:487-496.

297. Abramczyk H, Brozek-Pluska B, Surmacki J, Jablonska-Gajewicz J, Kordek R. Hydrogen bonds of interfacial water in human breast cancer tissue compared to lipid and DNA interfaces. *Journal of Biophysical Chemistry*. 2011;2:159.

298. Surmacki J, Musial J, Kordek R, Abramczyk H. Raman imaging at biological interfaces: applications in breast cancer diagnosis. *Mol Cancer*. 2013;12:48.

299. Liao Z, Lizio MG, Corden C, Khout H, Rakha E, Notingher I. Feasibility of integrated high-wavenumber Raman imaging and fingerprint Raman spectroscopy for fast margin assessment in breast cancer surgery. *J Raman Spectrosc*. 2020.

300. Stone N, Matousek P. Advanced transmission Raman spectroscopy: a promising tool for breast disease diagnosis. *Cancer Res*. 2008;68:4424-4430.

301. Stone N, Baker R, Rogers K, Parker AW, Matousek P. Subsurface probing of calcifications with spatially offset Raman spectroscopy (SORS): future possibilities for the diagnosis of breast cancer. *Analyst*. 2007;132:899-905.

302. Merritt S, Gulsen G, Chiou G, Chu Y, Deng C, Cerussi AE*, et al.* Comparison of water and lipid content measurements using diffuse optical spectroscopy and MRI in emulsion phantoms. *Technol Cancer Res Treat*. 2003;2:563-569.

303. Bowman T, Walter A, Shenderova O, Nunn N, McGuire G, El-Shenawee M. A phantom study of terahertz spectroscopy and imaging of micro-and nano-diamonds and nano-onions as contrast agents for breast cancer. *Biomed Phys Eng Express*. 2017;3:055001.

304. Masson LE, O'Brien CM, Pence IJ, Herington JL, Reese J, van Leeuwen TG*, et al.* Dual excitation wavelength system for combined fingerprint and high wavenumber Raman spectroscopy. *Analyst*. 2018;143:6049-6060.

305. Fowler P, Casey C, Cameron G, Foster MA, Knight C. Cyclic changes in composition and volume of the breast during the menstrual cycle, measured by magnetic resonance imaging. *BJOG*. 1990;97:595-602.

306. Tromberg BJ, Cerussi A, Shah N, Compton M, Durkin A, Hsiang D*, et al.* Imaging in breast cancer: diffuse optics in breast cancer: detecting tumors in pre-menopausal women and monitoring neoadjuvant chemotherapy. *Breast Cancer Res*. 2005;7:279-285.

307. Balleyguier C, Ayadi S, Van Nguyen K, Vanel D, Dromain C, Sigal R. BIRADS™ classification in mammography. *Eur J Radiol*. 2007;61:192-194.

308. Vardaki MZ, Gardner B, Stone N, Matousek P. Studying the distribution of deep Raman spectroscopy signals using liquid tissue phantoms with varying optical properties. *Analyst*. 2015;140:5112-5119.

309. Vardaki MZ, Matousek P, Stone N. Characterisation of signal enhancements achieved when utilizing a photon diode in deep Raman spectroscopy of tissue. *Biomed Opt Express*. 2016;7:2130-2141.

310. Hodson L, Skeaff CM, Fielding BA. Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake. *Prog Lipid Res*. 2008;47:348-380.

311. Michels R, Foschum F, Kienle A. Optical properties of fat emulsions. *Optics Express*. 2008;16:5907-5925.

312. Kalyanam S, Yapp RD, Insana MF. Poro-viscoelastic behavior of gelatin hydrogels under compression-implications for bioelasticity imaging. *J Biomech Eng*. 2009;131:081005.

313. Baxter Hospital Products. IV Fat emulsions Intralipid 20% [Internet]. Baxter medication; 2019 [Accessed 04/07/2019]. Available from: http://www.baxtermedicationdeliveryproducts.com/nutrition/ivfatemulsions.html.

314. Jacques SL. Optical properties of biological tissues: a review. *Phys Med Bi*. 2013;58:R37.

315. Tydeman C. The further assessment of a handheld Raman Spectrsocopy probe for the intraoperative diagnosis of axillary lymph nodes in breast cancer. online repositry: Cranfield University; 2012.

316. Ghita A, Matousek P, Stone N. Sensitivity of Transmission Raman Spectroscopy Signals to Temperature of Biological Tissues. *Sci Rep*. 2018;8:8379.

317. Adank MW, Fleischer JC, Dankelman J, Hendriks BHW. Real-time oncological guidance using diffuse reflectance spectroscopy in electrosurgery: the effect of coagulation on tissue discrimination. *J Biomed Opt*. 2018;23:1-10.

318. Guerbet. Patent Blue V [Internet]. Guerbet; 2019 [Accessed 09/07/2019]. Available from: http://guerbet-womenhealthcare.com/patent-blue-v/.

319. Laucirica R. Intraoperative assessment of the breast: guidelines and potential pitfalls. *Arch Pathol Lab Med*. 2005;129:1565-1574.

320. Hobro AJ, Konishi A, Coban C, Smith NI. Raman spectroscopic analysis of malaria disease progression via blood and plasma samples. *Analyst*. 2013;138:3927-3933.

321. Adar F. Molecular Spectroscopy Workbench Use of Raman Spectroscopy and 2D-COS for the Detection and Characterization of Chemical Interactions [Internet]. Spectroscopyonline; 2019 [Accessed 23/09/2020]. Available from: https://www.spectroscopyonline.com/view/use-raman-spectroscopy-and-2d-cos-detection-and-characterization-chemical-interactions.

322. Burikov S, Dolenko T, Patsaeva S, Starokurov Y, Yuzhakov V. Raman and IR spectroscopy research on hydrogen bonding in water–ethanol systems. *Mol Phys*. 2010;108:2427-2436.

323. McCreery RL. Chapter 10 Calibration and Validation. In: Winefordner JD, editor. Raman Spectroscopy for Chemical Analysis. 1st ed. New York: John Wiley & Sons; 2000. p. 251-291.

324. Liu D, Byrne HJ, O'Neill L, Hennelly B, editors. Investigation of wavenumber calibration for Raman spectroscopy using a polymer reference. Optical Sensing and Detection V; 2018: International Society for Optics and Photonics.

325. Frost KJ, McCreery RL. Calibration of Raman spectrometer instrument response function with luminescence standards: an update. *Appl Spectrosc*. 1998;52:1614-1618.

326. Rodriguez JD, Westenberger BJ, Buhse LF, Kauffman JF. Standardization of Raman spectra for transfer of spectral libraries across different instruments. *Analyst*. 2011;136:4232-4240.

327. Hutsebaut D, Vandenabeele P, Moens L. Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst*. 2005;130:1204-1214.

328. Ocean Optics Inc. Calibrating the Wavelength of the Spectrometer [Internet]. Ocean Optics Inc; 2019 [Accessed 30/07/2019]. Available from: https://oceanoptics.com/wp-content/uploads/Spectrometer-Wavelength-Calibration-Instructions.pdf.

329. Wollman S, Bohn P. Evaluation of polynomial fitting functions for use with CCD arrays in Raman spectroscopy. *Appl Spectrosc*. 1993;47:125-126.

330. Chen H, Zhang ZM, Miao L, Zhan DJ, Zheng YB, Liu Y*, et al.* Automatic standardization method for Raman spectrometers with applications to pharmaceuticals. *J Raman Spectrosc*. 2015;46:147-154.

331. McCreery RM. Standard Spectra [Internet]. University of Alberta; 2014 [Accessed 30/07/2019]. Available from: https://www.chem.ualberta.ca/~mccreery/ramanmaterials.html.

332. Gniadecka M, Nielsen OF, Christensen DH, Wulf HC. Structure of water, proteins, and lipids in intact human skin, hair, and nail. *J Invest Dermatol*. 1998;110:393-398.

333. Leikin S, Parsegian V, Yang W-H, Walrafen G. Raman spectral evidence for hydration forces between collagen triple helices. *Proc Natl Acad Sci U S A*. 1997;94:11312-11317.

334. Unal M, Yang S, Akkus O. Molecular spectroscopic identification of the water compartments in bone. *Bone*. 2014;67:228-236.

335. Bauer N, Wicksted JP, Jongsma F, March WF, Hendrikse F, Motamedi M. Noninvasive assessment of the hydration gradient across the cornea using confocal Raman spectroscopy. *Investig Ophthalmol Vis Sci*. 1998;39:831-835.

336. Abramczyk H, Brozek-Pluska B, Krzesniak M, Kopec M, Morawiec-Sztandera A. The cellular environment of cancerous human tissue. Interfacial and dangling water as a “hydration fingerprint”. *Spectrochim Acta A Mol Biomol Spectrosc*. 2014;129:609-623.

337. Barroso E, Bakker Schut T, Caspers P, Santos IP, Wolvius E, Koljenović S*, et al.* Characterization and subtraction of luminescence background signals in high-wavenumber Raman spectra of human tissue. *J Raman Spectrosc*. 2018;49:699-709.

338. Teledyne Princeton Instruments. PIXIS cameras [Internet]. Teledyne Princeton Instruments,; 2019 [Accessed 02/10/2019]. Available from: https://www.princetoninstruments.com/userfiles/files/assetLibrary/Datasheets/Princeton_Instruments_PIXIS_1024_eXcelon_N3_1_10_27_14.pdf.

339. Oxford Instruments. Andor iDus 1.7um InGaAs Specification sheet [Internet]. Oxford Instruments; 2019 [Accessed 02/10/2019]. Available from: https://andor.oxinst.com/products/idus-spectroscopy-cameras/idus-1-7-ingaas.

340. McCreery RL. Chapter 9 Nondispersive Raman spectrometers. In: JD W, editor. Raman Spectroscopy for Chemical Analysis. New York: John Wiley & Sons; 2000. p. 221-251.

341. Brindha E, Rajasekaran R, Aruna P, Koteeswaran D, Ganesan S. High wavenumber Raman spectroscopy in the characterization of urinary metabolites of normal subjects, oral premalignant and malignant patients. *Spectrochim Acta A Mol Biomol Spectrosc*. 2017;171:52-59.

342. McCreery RL. Chapter 4 Signal- To - Noise in Raman Spectroscopy. In: Winefordner JD, editor. Raman Spectroscopy for Chemical Analysis. 1st ed. New York: John Wiley & Sons; 2000. p. 49-73.

343. de Bree E, Mamalakis G, Sanidas E, Hatzis C, Askoxylakis I, Daskalakis M*, et al.* Adipose tissue fatty acid composition in Greek patients with breast cancer versus those with benign breast tumors. *Anticancer Res*. 2013;33:1667-1672.

344. Vlachos N, Skopelitis Y, Psaroudaki M, Konstantinidou V, Chatzilazarou A, Tegou E. Applications of Fourier transform-infrared spectroscopy to edible oils. *Anal Chim Acta*. 2006;573:459-465.

345. Yang S, Sen C, Zhou J-G, editors. InGaAs detector based Raman spectroscopy for water probing in biological tissue. Biophysics, Biology and Biophotonics IV: the Crossroads; 2019: International Society for Optics and Photonics.

346. Cerussi AE, Berger AJ, Bevilacqua F, Shah N, Jakubowski D, Butler J*, et al.* Sources of absorption and scattering contrast for near-infrared optical mammography. *Acad Radiol*. 2001;8:211-218.

347. Sandell JL, Zhu TC. A review of in-vivo optical properties of human tissues and its impact on PDT. *J Biophotonics*. 2011;4:773-787.

348. Kim S, Byun KM, Lee SY. Influence of water content on Raman spectroscopy characterization of skin sample. *Biomed Opt Express*. 2017;8:1130-1138.

349. Rylander CG, Stumpp OF, Milner TE, Kemp NJ, Mendenhall JM, Diller KR*, et al.* Dehydration mechanism of optical clearing in tissue. *J Biomed Opt*. 2006;11:041117.

350. Krishnakumar V, Prabavathi N. Simulation of IR and Raman spectral based on scaled DFT force fields: a case study of 2-amino 4-hydroxy 6-trifluoromethylpyrimidine, with emphasis on band assignment. *Spectrochim Acta A Mol Biomol Spectrosc*. 2008;71:449-457.

351. Unal M, Akkus O. Raman spectral classification of mineral-and collagen-bound water's associations to elastic and post-yield mechanical properties of cortical bone. *Bone*. 2015;81:315-326.

352. Li Q, Hao C, Xu Z. Diagnosis of breast cancer tissues using 785 nm miniature Raman spectrometer and pattern regression. *Sensors*. 2017;17:627.

353. Little MA, Varoquaux G, Saeb S, Lonini L, Jayaraman A, Mohr DC*, et al.* Using and understanding cross-validation strategies. Perspectives on Saeb et al. *Gigascience*. 2017;6:gix020.

354. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Statistics surveys*. 2010;4:40-79.

355. Saeb S, Lonini L, Jayaraman A, Mohr DC, Kording KP. The need to approximate the use-case in clinical machine learning. *Gigascience*. 2017;6:gix019.

356. Lazaro-Pacheco D, Shaaban AM, Rehman S, Rehman I. Raman spectroscopy of breast cancer. *Appl Spectrosc Rev*. 2019:1-37.

357. Kumar V CR, Robbins S. Female Genital system and Breast. In: Robbins P, editor. Basic Pathology. 6 ed. Philadelphia: W.B.Saunders; 1997. p. 597-637.

358. Thakur SB, Brennan SB, Ishill NM, Morris EA, Liberman L, Dershaw DD*, et al.* Diagnostic usefulness of water-to-fat ratio and choline concentration in malignant and benign breast lesions and normal breast parenchyma: an in vivo (1) H MRS study. *J Magn Reson Imaging*. 2011;33:855-863.

359. Lawrence G KO, Lagord C, Cheung S, Sidhu J, Sagar C. Second all breast cancer report. Web: National Cancer Intelligence Network,; 2011 2011.

360. Hughes L, Mansel R, Webster DT. Aberrations of normal development and involution (ANDI): a new perspective on pathogenesis and nomenclature of benign breast disorders. *Lancet*. 1987;330:1316-1319.

361. Brem RF, Ioffe M, Rapelyea JA, Yost KG, Weigert JM, Bertrand ML*, et al.* Invasive lobular carcinoma: detection with mammography, sonography, MRI, and breast-specific gamma imaging. *AJR Am J Roentgenol*. 2009;192:379-383.

362. Ewen Smith GD. Chapter 2: The Raman Experiment. In: Wiley, editor. Modern Raman Spectroscopy. 1st ed. Chichester: John Wiley & Sons Ltd; 2005.

363. Santos IsP, Caspers PJ, Bakker Schut TC, van Doorn R, Noordhoek Hegt V, Koljenović S*, et al.* Raman spectroscopic characterization of melanoma and benign melanocytic lesions suspected of melanoma using high-wavenumber Raman spectroscopy. *Anal Chem*. 2016;88:7683-7688.

364. Czamara K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M. Raman spectroscopy of lipids: a review. *J Raman Spectrosc*. 2015;46:4-20.

365. Santos lP, Caspers PJ, Bakker Schut T, van Doorn R, Koljenović S, Puppels GJ. Implementation of a novel low-noise InGaAs detector enabling rapid near-infrared multichannel Raman spectroscopy of pigmented biological samples. *J Raman Spectrosc*. 2015;46:652-660.

366. Kong K, Kendall C, Stone N, Notingher I. Raman spectroscopy for medical diagnostics--From in-vitro biofluid assays to in-vivo cancer detection. *Adv Drug Deliv Rev*. 2015;89:121-134.

367. Matousek P. Deep non-invasive Raman spectroscopy of living tissue and powders. *Chem Soc Rev*. 2007;36:1292-1304.

368. Keller MD, Majumder SK, Mahadevan-Jansen A. Spatially offset Raman spectroscopy of layered soft tissues. *Opt Lett*. 2009;34:926-928.

369. Pappo I, Spector R, Schindel A, Morgenstern S, Sandbank J, Leider LT*, et al.* Diagnostic performance of a novel device for real-time margin assessment in lumpectomy specimens. *J Surg Res*. 2010;160:277-281.

370. Thill M, Dittmer C, Baumann K, Friedrichs K, Blohmer J-U. MarginProbe®–Final results of the German post-market study in breast conserving surgery of ductal carcinoma in situ. *Breast J*. 2014;23:94-96.

371. Ghita A, Hubbard T, Matousek P, Stone N. Non-invasive Detection of Differential Water Content inside Biological Samples using Deep Raman Spectroscopy. *Anal Chem*. 2020;92:9449-9453.

372. Grant Y, Al-Khudairi R, St John E, Barschkett M, Cunningham D, Al-Mufti R*, et al.* Patient-level costs in margin re-excision for breast-conserving surgery. *Br J Surg*. 2019;106:384-394.

373. Cochrane R, Valasiadou P, Wilson A, Al-Ghazal S, Macmillan R. Cosmesis and satisfaction after breast-conserving surgery correlates with the percentage of breast volume excised. *Br J Surg*. 2003;90:1505-1509.

374. Rivera RJ, Holmes DR, Tafra L. Analysis of the impact of intraoperative margin assessment with adjunctive use of MarginProbe versus standard of care on tissue volume removed. *Int J Surg Oncol*. 2012;2012.

375. Bundred NJ. A trial looking at reducing further operations after breast conserving surgery [Internet]. Cancer Research UK; 2019 [Accessed 05/06/2020]. Available from: https://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-trial-looking-at-reducing-further-operations-after-breast-conserving-surgery#undefined.

376. St John ER, Balog J, McKenzie JS, Rossi M, Covington A, Muirhead L*, et al.* Rapid evaporative ionisation mass spectrometry of electrosurgical vapours for the identification of breast pathology: towards an intelligent knife for breast cancer surgery. *Breast Cancer Res*. 2017;19:59.

377. Goldberg BD, Iftimia NV, Bressner JE, Pitman MB, Halpern EF, Bouma BE*, et al.* Automated algorithm for differentiation of human breast tissue using low coherence interferometry for fine needle aspiration biopsy guidance. *J Biomed Opt*. 2008;13:014014.