The iKnife: Real Time Intra-Operative Tissue Diagnostics using Rapid Evaporative Ionisation Mass Spectrometry

Thesis submitted for the degree of Doctor of Philosophy

Miss Laura Jane Muirhead MBChB MRCS.
Department of Surgery and Cancer
Imperial College, London

Supervisors: Professor A Darzi, Professor J K Nicholson, Professor Z Takats,
Dr J Kinross
For my parents,
Who never question, only believe.

And for my sister,
Who is always there.
Abstract

Background: This thesis describes the development of a novel technique for intra-operative tissue diagnostics, which aims to provide the operating surgeon with molecular diagnostic information in vivo and in near real time by the application of mass spectrometry analysis to tissue specific ions contained within the metabolite rich electrosurgery smoke plume. The system is known as Rapid Evaporative Ionisation Mass Spectrometry (REIMS) or the ‘iKnife’ and it utilises recent developments in systems biology and spectroscopic technology to provide accurate, real time metabolic information in the clinical and operating theatre environment. In addition to tissue diagnostics, REIMS is a potential tool for investigation of the biological processes underpinning the detected lipid changes, contributing to a greater understanding of tumour biology and lipid profiling and identification of novel prognostic or therapeutic targets.

Methodology: This body of work developed the instrumentation and tested and validated the analytical methodology for REIMS. Prospective observational studies of breast (N= 45) and colorectal cancer (N=40) were performed at Imperial College NHS trust, for the identification of histologically discrete specific lipidomic profiles. Finally, an endoscopic application was created for the translational application of this technology in the diagnosis and treatment of colorectal cancer.

Results: Optimal machine set up was obtained using a Waters Xevo G2-S QTof i-Knife mass spectrometer with a modified atmospheric interface. Multivariate analysis lead to the identification of colorectal adenocarcinoma and invasive breast cancer models through the quantitative analysis of lipid profiles. Significant lipid species responsible for the distinction of healthy tissue and cancer were identified and, based on these profiles, colorectal and breast cancer were identified with a diagnostic accuracy of 93.6% and 88.5% respectively. In addition, tumour characteristics including grade, stage, presence of genetic mutations and hormone receptor status influenced spectral profiles and allowed further discrimination of tumour samples. An endoscopic tool for near real time analysis of colorectal tissue accurately identified the anatomical layers of the colonic wall and discriminated colorectal
adenocarcinoma from normal mucosa, ex vivo, with a diagnostic accuracy of 88% and 89.5% respectively. This work lead to the development of fully functioning iKnife and iendoscope instrumental platforms

**Conclusions:** The REIMS platform developed in this thesis is a fundamental first step in facilitating future in vivo REIMS experimentation to determine its potential value as a clinical tool. This analysis has optimized, standardized and validated a methodology for breast and colorectal tissue sampling, and it has identified candidate lipid biomarkers of translational use.
Declaration

This thesis, and material presented is the result of my own work. Preparation of blocks for histological analysis was carried out by Ms Anna Mroz (Department of Surgery and Cancer, Imperial College, London). The Matlab toolbox used for histological co-registration and multivariate analysis of REIMS and DESI data was designed and built by Dr Kirill Veselkov (Department of Computational and Systems Medicine, Imperial College, London) and Matlab scripts for data analysis provided by Mr Ottmar Golf (Department of Computational and Systems Medicine, Imperial College, London) The DESI-REIMS cross validation algorithm was developed by Dr Kirill Veselkov and Mr Ottmar Golf. The interface for DESI tissue analysis was built and tested by Ms Nicole Strittmatter and Professor Zoltan Takats.

Conflicts of Interest and Funding

Several grants were awarded to fund completion of this work. They are as follows:


• European Commission FP7 Programme Grant. The Intelligent Surgical Device. (£2,443,474.60). Participants 1. Medimass (Hungary) 2. Imperial College London (UK) 3. Asklepios Medical School (Germany). Start date 01/07/2012. Proposed end date 31/08/2015. Project closed 31/10/2013.

The project was closed in 2013 in order to evaluate business opportunities with potential investors and
partners, due to commercial interest in the device under development by MediMass Ltd, in conjunction with Imperial College London.

For the preceding 3 years our research group at Imperial College London had been working in close partnership with Micromass UK Ltd, a fully owned subsidiary of the large commercial mass spectrometer manufacturer Waters Corporation, in development of the mass spectrometry instrumentation required for REIMS tissue analysis. In July 2014 Waters acquired REIMS technology; including patent applications, software, databases and REIMS expertise, from MediMass Ltd. Imperial researchers continue to work with Waters scientists to further develop the technology. Professor A Darzi, Professor Z Takats and Professor J K Nicholson receive consultancy fees from Waters Corporation. Ongoing research and development work continues to be carried out within Imperial College London in collaboration with Waters.

As part of the acquisition process, in November 2015 I conducted 2 telephone consultations with Boston Strategic Partners, inc. consulting on Behalf of Waters Corporation, in an advisory capacity to map out the REIMS pathology and database development workflow, for which I was financially remunerated.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives license. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work.
Acknowledgments

Firstly, I would like to thank my supervisors Professor Ara Darzi (Biosurgery & Surgical Technology, Imperial College London), Professor Zoltan Takats (Computational & Systems Medicine, Imperial College London), Professor Jeremy Nicholson (Computational & Systems Medicine, Imperial College London) and Dr James Kinross (Biosurgery & Surgical Technology, Imperial College London) for their patience, expert guidance and encouragement. The outstanding combined level of knowledge and expertise they shared with me during my time under their supervision was truly extraordinary, and will continue to inspire me throughout future endeavors.

I also owe a great deal of thanks to my colleagues in the Histopathology Department (Centre for Pathology, Department of Medicine, Imperial College London). Dr Abigail Speller for her invaluable input to this work, assisting in safe tissue sample collection, histological examination of all tissue samples and providing advice and support throughout. Professor Rob Goldin and Dr Mike Osborn for sharing their expert knowledge and opinion on GI tract pathology and providing samples of GI tract tissue for this research. Professor Sami Shousha and Dr Rathi Ramkrishnan for providing samples of breast tissue from surgical resections and their expert knowledge of breast pathology. Ms Anna Mroz for all of her assistance with tissue preparation.

I am exceptionally grateful to my colleagues in the Department of Computational and Systems Medicine, Imperial College London for their support in scientific pursuits. Dr Julia Balog for her advice and support in all things surrounding mass spectrometry instrumentation and data collection. Dr Kirill Veselkov and Mr Ottmar Golf for developing the data analysis tools required to perform this work, and their assistance with data analysis. Ms Nicole Strittmatter for construction of the DESI MSI interface and for her teaching and supervision in DESI analysis. Dr Emrys Jones and Dr Steve Pringle for sharing their expert knowledge of mass spectrometry instrumentation and contribution to the development of the REIMS capable mass spectrometers and interfaces in this work.

Finally I would like to thank my surgical colleagues within Imperial College Healthcare NHS Trust. Mr Barry Paraskeva, Mr Paul Ziprin, Professor Dimitri Hadjiminas, Mrs Katy Hogben, Mrs Jaqueline Lewis
and the operating theatre staff for their enthusiastic co-operation and assistance with patient recruitment, tissue collection and clinical data collection.
9 Appendices .................................................................................................................................................. 217
List of Tables

TABLE 1: SUMMARY OF CURRENT AND POTENTIAL FUTURE BIOMARKERS FOR USE IN THE SCREENING, DIAGNOSIS, PROGNOSTICATION AND CLINICAL MANAGEMENT OF COLORECTAL CANCER.... 42

TABLE 2: BREAST CANCER SUBTYPES ........................................................................................................... 45

TABLE 3: LEVEL OF HER2 PROTEIN EXPRESSION AND CORRESPONDING CLINICAL LABORATORY RESULT ............................................................................................................................. 48

TABLE 4: MAS NMR/1H NMR STUDIES IN BREAST CANCER TISSUE/EXTRACT DEMONSTRATING METABOLITE ALTERATIONS IN CANCER TISSUE COMPARED TO HEALTHY TISSUE AND MEASURABLE METABOLITE CHANGES ASSOCIATED WITH INDIVIDUAL ONCOLOGICAL PROGNOSTIC FEATURES. 53

TABLE 5: MAS NMR/1H NMR STUDIES IN COLORECTAL CANCER TISSUE/EXTRACT DEMONSTRATING METABOLITE ALTERATIONS IN CANCER TISSUE COMPARED TO HEALTHY TISSUE AND MEASURABLE METABOLITE CHANGES ASSOCIATED WITH ONCOLOGICAL CLINICAL FEATURES .............................................. 55

TABLE 6: CURRENT TECHNIQUES FOR USE IN THE INTRA-OPERATIVE ASSESSMENT OF BREAST CANCER RESECTION MARGINS .......................................................................................................................... 62

TABLE 7: COMMONLY UTILISED MASS SPECTROMETRY PLATFORMS FOR THE ANALYSIS OF HUMAN BIOFLUIDS AND TISSUE ............................................................................................................. 68

TABLE 8: CLINICOPATHOLOGICAL CHARACTERISTICS OF COLORECTAL TISSUE SAMPLES USED IN TESTING OF THE WATERS SYNAPT G2-S ................................................................. 106

TABLE 9: CLINICOPATHOLOGICAL DETAILS OF COLORECTAL SAMPLES USED TO TEST THE XEVO G2-S QTOF I-KNIFE SYSTEM ................................................................................................................................. 111

TABLE 10: CLINICOPATHOLOGICAL CHARACTERISTICS OF COLORECTAL TISSUE SAMPLES USED TO BUILD REIMS DATABASE ......................................................................................................... 118

TABLE 11: MISCLASSIFIED SAMPLES OF COLORECTAL ADENOCARCINOMA AND NORMAL HEALTHY MUCOSA ............................................................................................................................ 122

TABLE 12: SUMMARY OF DYSPLASTIC COLORECTAL SAMPLES ..................................................................................... 130

TABLE 13: PEAKS MOST HIGHLY SIGNIFICANT FOR THE DIFFERENTIATION OF CANCER AND NORMAL ADJACENT MUCOSA. THE EXACT MASS AND GLYCEROPHOSPHOLIPID TYPE OF EACH SPECIES ARE SPECIFIED BELOW THE STRUCTURE, WITH THE NUMBER OF CARBON ATOMS AND DOUBLE BONDS IN PARENTHESES, AND AN INDICATION OF EITHER LOSS OF HYDROGEN OR AMMONIA INDICATING HOW THE NEGATIVE ION IS GENERATED ................................................................................................. 133

TABLE 14: CLINICAL CHARACTERISTICS OF BREAST TISSUE SAMPLES .............................................................................. 144

TABLE 15: HISTOLOGICAL CLASSIFICATION .......................................................................................................................... 145

TABLE 16: BREAST CANCER SAMPLES MISCLASSIFIED AS NORMAL HEALTHY BREAST TISSUE .................................. 148

TABLE 17: MISCLASSIFIED BREAST CANCER SAMPLES ........................................................................................................... 151
TABLE 18: MOST STATISTICALLY SIGNIFICANT M/Z PEAKS RESPONSIBLE FOR THE DIFFERENTIATION OF NORMAL AND MALIGNANT BREAST TISSUE, THE LIPID SPECIES, CHEMICAL STRUCTURE AND P VALUE .................................................................151

TABLE 19: DESI AND REIMS TISSUE SAMPLES USED IN THE ANALYSIS .................................................................162

TABLE 20: STATISTICALLY SIGNIFICANT PEAKS IN THE DESI ANALYSIS OF COLORECTAL CANCER WITH LIPID SPECIE AND P VALUE .................................................................167

TABLE 21: CLINICOPATHOLOGICAL CHARACTERISTICS OF TISSUE SAMPLES USED IN THE ENDOSCOPIC REIMS ANALYSIS OF COLORECTAL ADENOCARCINOMA AND NORMAL ADJACENT MUCOSA. BS REPRESENTS THE NUMBER OF DATAPoints PRODUCED FROM THE REIMS ANALYSIS OF EACH SAMPLE. THIS IS DEPENDANT ON THE SIZE OF THE SAMPLE WITH LARGER SAMPLES PROVIDING MORE REPLICATES ..............................................................................175

TABLE 22: SAMPLES MISCLASSIFIED IN IENDOSCOPE ANALYSIS ..............................................................................177
List of Figures

FIGURE 1: A) HALSTED RADICAL MASTECTOMY DEMONSTRATING REMOVAL OF THE BREAST, PECTORAL MUSCLES, AXILLARY LYMPHNODES AND SKIN (IMAGE FROM HALSTED. THE RESULTS OF OPERATIONS FOR THE CURE OF CANCER OF THE BREAST PERFORMED AT THE JOHNS HOPKINS HOSPITAL FROM JUNE, 1889, TO JANUARY, 1894 ANN SURG. 1894 NOV; 20(5); 497-555) AND B) PICTORIAL REPRESENTATION OF A MODERN WIDE LOCAL EXCISION OF BREAST CANCER. 21

FIGURE 2: A) PRINCIPAL COMPONENTS ANALYSIS (UNSUPERVISED) OF A DATASET WITH 5 INDIVIDUAL CLASSES. MAXIMAL VARIANCE WITHIN THE DATA CAN BE SEEN ALONG PCI. B) LINEAR DISCRIMINANT ANALYSIS (SUPERVISED) IDENTIFIES THE AXES RESPONSIBLE FOR THE GREATEST DIFFERENTIATION BETWEEN CLASSES AND HAS MAXIMISED SEPARATION BETWEEN THE GROUPS. 30

FIGURE 3: EXAMPLE MASS SPECTRUM (M/Z RANGE 300-640) DEMONSTRATING M/Z VALUE ON THE X-AXIS AND RELATIVE ABUNDANCE OF PEAKS ON THE Y-AXIS. 69

FIGURE 4: 1. OVERVIEW OF REIMS INSTRUMENTATION AND DATA COLLECTION USING MONOPOLAR ELECTROSURGERY AS THE IONISATION METHOD. A/B) EXAMPLE REIMS MASS SPECTRA OF HUMAN TISSUE IN THE PHOSPHOLIPID M/Z RANGE (600-900 M/Z) WITH SPECIFIC PHOSPHOLIPID PEAKS HIGHLIGHTED. C) PREDOMINANT DISCRIMINATORY MEMBRANE PHOSPHOLIPID SPECIES DETECTED BY REIMS. IMAGE MODIFIED FROM BALOG ET AL INTRAOPERATIVE TISSUE IDENTIFICATION USING RAPID EVAPORATIVE IONIZATION MASS SPECTROMETRY, SCI TRANSL MED 5, 194RA93 (2013). 71

FIGURE 5: SCHEMATIC DIAGRAM OF DESI INSTRUMENTAL SETUP WITH MOBILE ELECTROSPRAY SOURCE, INLET CAPILLARY AND SAMPLE STAGE. B) IONISATION OF THE TISSUE SAMPLE SURFACE OCCURS IN A RASTER PATTERN. EACH MOVEMENT OF THE SAMPLE UNDER THE SPRAYER GENERATES ITS OWN, TOPOGRAPHICALLY ACCURATE, MASS SPECTRUM, REPRESENTING THE CHEMICAL COMPOSITION OF THE SAMPLE SURFACE IN THAT AREA. USING THIS INFORMATION, A CHEMICAL IMAGE OF THE SAMPLE IS FORMED, BASED ON THE MS PROFILE DERIVED FROM EACH SAMPLING POINT. FIGURE MODIFIED WITH PERMISSION FROM MASS SPECTROMETRY SAMPLING UNDER AMBIENT CONDITIONS WITH DESORPTION ELECTROSPRAY IONISATION. TAKATS ET AL. SCIENCE. 2004 OCT 15;306(5695):471-3 75

FIGURE 6: EXAMPLE REIMS SPECTRAL DATA FILES OBTAINED FROM EX VIVO SAMPLING OF COLORECTAL ADENOCARCINOMA AND NORMAL ADJACENT MUCOSA WITH THE PHOSPHOLIPID REGION OF INTEREST (600-900 M/Z) HIGHLIGHTED. MAGNIFIED VIEWS OF THE CORRESPONDING H&E STAINED SLIDE OF EACH TISSUE SAMPLE ARE SHOWN, WITH VISIBLE ELECTROSURGERY CUTS, USED FOR HISTOLOGICAL CLASSIFICATION OF THE SPECTRAL DATA FILE WITHIN THE DATABASE. 76

FIGURE 7: ALGORITHM FOR REIMS DATABASE MODELLING AND UNKNOWN TISSUE IDENTIFICATION. 76

FIGURE 8: EXAMPLE CONFUSION MATRIX. 77

FIGURE 9: EXAMPLE ROC CURVE WITH HIGH AUC REPRESENTING AN ACCURATE DISCRIMINATORY TEST. 80

FIGURE 10: SAMPLE OF BREAST TISSUE DEMONSTRATING 4 DISTINCT NEEDLE ELECTRODE SAMPLING POINTS AND THE CORRESPONDING MICROSCOPIC IMAGE OF NORMAL FIBRO-FATTY BREAST TISSUE WITH A VISIBLE ELECTROSURGERY SAMPLING POINT USED TO HISTOLOGICALLY CLASSIFY THE MS DATAFILE WITHIN THE DATABASE. 94

FIGURE 11: A) SCHEMATIC DIAGRAM OF THE ELECTROSURGICAL SNARE SYSTEM WITHIN THE COLONOSCOPE, CONNECTED TO THE Mass SPECTROMETER FOR SMOKE PLUME ANALYSIS. B) MODIFIED ENDOSCOPIC SNARE WITH PUNCTURE HOLES IN THE PLASTIC INSULATOR AND THE ADDITION OF PTFE TUBING VIA A THREE-WAY STOPCOCK VALVE DISTAL TO THE SNARE HANDLE. 95

14
FIGURE 12: DESI DATA ANALYSIS WORKFLOW. A) SLIDE SCANNED IMAGE OF TISSUE SECTION WITH ADENOCARCINOMA, HEALTHY MUCOSA AND STROMA ANNOTATED BY HISTOPATHOLOGIST. B) CORRESPONDING DESI IMAGE WITH EACH TISSUE TYPE NOW IDENTIFIED BASED ON SPECTRAL DATA. C) RAW DESI SPECTRAL DATA REPRESENTING EACH TISSUE TYPE. D) SUPERVISED ANALYSIS DEMONSTRATING DIFFERENTIATION OF EACH TISSUE TYPE BASED ON DESI MASS SPECTRAL DATA

FIGURE 13: MODIFIED REIMS MONOPOLAR ELECTROSURGERY HANDPIECE

FIGURE 14: CUSTOMISED WATERS SYNAPT G2-S

FIGURE 15: A/B) RAW SPECTRAL OBTAINED FROM ANALYSIS OF BENIGN AND MALIGNANT COLORECTAL TISSUE ON THE WATERS SYNAPT G2-S DEMONSTRATING HIGH LEVELS OF BACKGROUND NOISE. C) AN EXAMPLE OF A GOOD QUALITY SPECTRUM, ALSO FROM ANALYSIS OF COLORECTAL MUCOSA, WITH HIGH SIGNAL TO NOISE RATIO AND IMPROVED SIGNAL INTENSITY

FIGURE 16: OVERVIEW OF ION OPTICS IN THE WATERS SYNAPT G2-S

FIGURE 17: WATERS XEVO G2-S QTOF I-KNIFE

FIGURE 18: WATERS XEVO G2-S I-KNIFE SYSTEM. A) CONFUSION MATRIX DEMONSTRATING POOR OVERALL CLASSIFICATION ACCURACY IN THE DIFFERENTIATION OF COLORECTAL ADENOCARCINOMA AND NORMAL ADJACENT MUCOSA USING XEVO G2-S I-KNIFE SYSTEM. B) PCA PLOT DEMONSTRATES VISUAL CLUSTERING OF SAMPLES (BLUE = HEALTHY MUCOSA, PINK = ADENOCARCINOMA) ANALYSED ON THE SAME DAY, THOUGHT TO BE DUE TO INSTRUMENT CONTAMINATION

FIGURE 19: ORIGINAL INLET CAPILLARY AND COLD BALL COLLISION SURFACE OF XEVO G2-S I-KNIFE DEMONSTRATING BUILD UP OF DEBRIS ON COLLISION SURFACE. NOVEL HEATED KANTHAL COIL INTERFACE OUTSIDE AND, HEATED, INSIDE THE INSTRUMENT

FIGURE 20: REIMS DATA FROM ANALYSIS OF PORCINE COLON MUCOSA SAMPLE USING ELECTROSURGERY POWER SETTINGS 30-70W

FIGURE 21: XEVO G2-S I-KNIFE INSTALLED FOR FUNCTIONALITY TESTING IN OPERATING THEATRE WITHIN ST MARY'S HOSPITAL, PADDINGTON

FIGURE 22: SUMMARY OF HEALTHY, BENIGN AND MALIGNANT COLORECTAL TISSUE SAMPLES COLLECTED AND SPECTRAL DATAPORTS OBTAINED

FIGURE 23: EXEMPLARY RAW SPECTRA FROM REIMS ANALYSIS OF NORMAL ADJACENT COLONIC MUCOSA AND COLORECTAL ADENOCARCINOMA DEMONSTRATING PHOSPHOLIPID PEAKS DETECTED IN THE 600-900 M/Z RANGE

FIGURE 24: A) PCA OF COLORECTAL ADENOCARCINOMA AND NORMAL ADJACENT MUCOSA B) LDA SCORES PLOT DEMONSTRATING MARKED SEPARATION OF THE 2 TISSUE TYPES C) CONFUSION MATRIX DEMONSTRATING OVERALL CLASSIFICATION ACCURACY OF 93%

FIGURE 25: HIGHLIGHTED MISCLASSIFIED SAMPLES OF NORMAL ADJACENT MUCOSA (GREEN WITH SURROUNDING RED CIRCLE) AND INVASIVE ADENOCARCINOMA (RED WITH GREEN CIRCLE) SAMPLE 44 S1 IS MOST SIGNIFICANTLY OUTLYING SAMPLE AND MAY HAVE BEEN MISLABELLED

FIGURE 26: A) PCA B) LDA OF COLONIC AND RECTAL ADENOCARCINOMA C) CONFUSION MATRIX DEMONSTRATING CLASSIFICATION OF SAMPLES BY LOCATION

FIGURE 27: LDA PLOTS, CONFUSION MATRICES AND ROC CURVES DEMONSTRATING PREDICTION OF T STAGE (T2 VS T3/T4), LYMPH NODE INVOLVEMENT AND LEVEL OF DIFFERENTIATION (MODERATELY DIFFERENTIATED VS POORLY DIFFERENTIATED) BY REIMS

FIGURE 28: LDA PLOTS, CONFUSION MATRICES AND ROC CURVES SHOWING MORE ACCURATE PREDICTION OF T STAGE (T2 VS T3/T4), LYMPH NODE INVOLVEMENT AND LEVEL OF DIFFERENTIATION BY REIMS IN RECTAL ADENOCARCINOMA ONLY

15
FIGURE 29: LDA PLOTS, CONFUSION MATRICES AND ROC CURVES DEMONSTRATING PREDICTION OF PRESENCE OF KRAS MUTATION AND EXTRA MURAL VASCULAR INVASION (EMVI) BY REIMS.

FIGURE 30: A) PCA AND B) LDA PLOTS OF COLORECTAL ADENOCARCINOMA, DYSPLASIA AND NORMAL ADJACENT MUCOSA C) CONFUSION MATRIX DEMONSTRATING OVERALL CLASSIFICATION ACCURACY OF THE MODEL FOR THE DISCRIMINATION OF ALL 3 TISSUE TYPES.

FIGURE 31: SCORES PLOTS, CONFUSION MATRICES AND ROC CURVES FROM ANALYSIS OF NAM AND DYSPLASIA, INVASIVE CANCER AND DYSPLASIA AND NAM AND ALL ABNORMAL TISSUE.

FIGURE 32: ANOVA WITH EXAMPLE BOX PLOTS OF STATISTICALLY SIGNIFICANT M/Z PEAKS, THE ASSOCIATED LIPID SPECIES AND P VALUE FOR THE DISCRIMINATION OF CANCER AND NAM.

FIGURE 33: REIMS RAW SPECTRAL DATA DEMONSTRATING FATTY ACID PEAKS AT M/Z 150-400, PHOSPHOLIPIDS AT 600-850 AND TRIGLYCERIDES FROM 850-1100.

FIGURE 34: SPECTRAL DATA OBTAINED FROM REIMS ANALYSIS OF LIPID RICH HEALTHY BREAST TISSUE, DEMONSTRATING A FEW FATTY ACID PEAKS ONLY, AND OF INVASIVE BREAST CANCER DEMONSTRATING TRIGLYCERIDE SIGNAL.

FIGURE 35: RAW SPECTRAL DATA ACQUIRED FROM PI REIMS ANALYSIS OF INVASIVE CANCER AND HEALTHY BREAST TISSUE DEMONSTRATING POOR SIGNAL ACQUISITION AND SOME TRIGLYCERIDE PEAKS IN INVASIVE CANCER TISSUE AND HIGH INTENSITY TRIGLYCERIDE PEAKS IN HEALTHY BREAST TISSUE.

FIGURE 36: ORIGINAL PI REIMS INTERFACE WITH FREESTANDING ELECTROSPRAY POSITIONED ORTHOGONALLY TO THE INLET CAPILLARY AND THE NEWLY DESIGNED INTERFACE WITH T PIECE PROVIDING DIRECT COUPLING OF ELECTROSPRAY AND INLET CAPILLARY.

FIGURE 37: EXAMPLE SPECTRAL PROFILES PRODUCED FROM THE ANALYSIS OF NORMAL BREAST TISSUE, BENIGN LESIONS AND INVASIVE CANCER USING THE FINAL IPA REIMS SETUP. SPECTRA ARE TYPICAL OF THOSE THAT WILL BE PRODUCED DURING ALL FURTHER ANALYSIS OF BREAST TISSUE.

FIGURE 38: SUMMARY OF BREAST TISSUE SAMPLES.

FIGURE 39: PCA AND LDA PLOTS OF BREAST CANCER TISSUE (RED) AND HEALTHY BREAST TISSUE (GREEN). CONFUSION MATRIX DEMONSTRATING CLASSIFICATION ACCURACY OF THE REIMS MODEL IN THE IDENTIFICATION OF BREAST CANCER TISSUE.

FIGURE 40: DATAPoints REPRESENTING BREAST CANCER CLUSTERING WITH NORMAL HEALTHY BREAST.

FIGURE 41: X10 (IN ADDITION TO STANDARD X10 EYEPIECE MAGNIFICATION) MAGNIFIED IMAGES OF "MIXED" BREAST TISSUE SAMPLES DEMONSTRATING RELATIONSHIP OF TUMOUR CELLS TO ELECTROSURGICAL BURN AND REIMS CLASSIFICATION. A) SMALL FOCUS OF ILC 0.1MM FROM ELECTROSURGICAL BURN POINT. THE RATIO OF ILC: FIBROUS TISSUE IN THE SAMPLE IS VERY LOW. B) SMALL FOCUS OF ILC 0.1MM FROM ELECTROSURGICAL BURN POINT. C) DENSE CANCER CELLS ON ONE SIDE OF THE ELECTROSURGICAL BURN AND ADIPOSE TISSUE (NO GLANDS) ON THE OTHER. BURN TRAVELS THROUGH BOTH TISSUE TYPES. D) FOCUS OF IDC 1.1MM FROM ELECTROSURGICAL BURN POINT.

FIGURE 42: LDA PLOT OF HEALTHY BREAST TISSUE (RED) AND MIXED SAMPLES (PURPLE) DEMONSTRATING 47 S1 AND 24 S1 AS OUTLIERS. 47 S1 IS IMAGE C IN FIGURE 44. THE ELECTROSURGICAL BURN TRAVELS THOUGH BOTH DENSE TUMOUR CELLS AND ADIPOSE TISSUE.

FIGURE 43: CONFUSION MATRIX AND ROC CURVE DEMONSTRATING CLASSIFICATION ACCURACY FOR THE IDENTIFICATION OF HEALTHY AND MALIGNANT BREAST TISSUE WITH MIXED SAMPLES REMOVED FROM THE DATASET.
FIGURE 44: ANOVA PLOT DEMONSTRATING PEAKS REPRESENTING THE MONOUNSATURATED PCS, PA, PI AND PE WHICH ARE STATISTICALLY SIGNIFICANT IN THE DISTINCTION OF BREAST CANCER AND HEALTHY BREAST TISSUE. .................................................................................................153

FIGURE 45: PCA AND LDA PLOTS OF HEALTHY BREAST TISSUE (GREEN), INVASIVE MALIGNANCY (RED) AND FIBROADENOMA (PURPLE). CONFOUND MATRIX SHOWING CLASSIFICATION ACCURACY FOR THE EX VIVO IDENTIFICATION OF HEALTHY BREAST TISSUE, INVASIVE CANCER AND FIBROADENOMA. ........................................................................................................154

FIGURE 46: CONFOUND MATRIX AND ROC CURVE DEMONSTRATING ACCURACY OF THE MODEL FOR THE CLASSIFICATION OF INVASIVE CANCER AND FIBROADENOMA (TOP) AND HEALTHY BREAST TISSUE AND FIBROADENOMA. ..................................................................................................................155

FIGURE 47: LDA SCORES PLOT AND ROC CURVE OF REIMS ANALYSIS OF BREAST IDC AND ILC.............................156

FIGURE 48: LDA PLOT, CONFOUND MATRIX AND ROC CURVE DEMONSTRATING ACCURACY IN THE DIFFERENTIATION OF HER2 POSITIVE AND NEGATIVE TUMOUR SAMPLES........................................................................157

FIGURE 49: LDA PLOT WITH CORRECTLY CLASSIFIED SAMPLES ILLUSTRATED, ROC CURVE AND CONFOUND MATRIX DEMONSTRATING OVERALL ACCURACY IN THE DISCRIMINATION OF ER POSITIVE AND NEGATIVE BREAST CANCER ...........................................................................................................159

FIGURE 50: LDA PLOT AND CONFOUND MATRIX PRODUCED BY REIMS ANALYSIS OF GRADE 1(GREEN), GRADE 2 (ORANGE) AND GRADE 3(RED) BREAST CANCER SAMPLES. ................................................160

FIGURE 51: DESI MSI ANALYSIS OF INDIVIDUAL TISSUE LAYERS OF NORMAL COLONIC WALL.........................164

FIGURE 52: DESI COLORECTAL ADENOCARCINOMA AND HEALTHY MUCOSA......................................................166

FIGURE 53: DESI ANALYSIS OF COLORECTAL ADENOCARCINOMA, DYSPLASIA AND NORMAL HEALTHY MUCOSA ........................................................................................................................................168

FIGURE 54: CONFOUND MATRIX DESI REIMS CROSS VALIDATION .......................................................................170

FIGURE 55: RAW SPECTRAL DATA PRODUCED DURING ANALYSIS OF PORCINE COLON USING THE REIMS ENDOSCOPIC SNARE SYSTEM ..................................................................................................................171

FIGURE 56: MAGNIFIED VIEW OF COLONIC WALL LAYERS AND CORRESPONDING RAW REIMS SPECTRAL DATA.........................................................................................................................................173

FIGURE 57: A) UNSUPERVISED (PCA) ANALYSIS DEMONSTRATING SEPARATION OF THE 3 COLONIC LAYERS. B) RAW MASS SPECTRA DEMONSTRATING SIGNIFICANT DIFFERENCES IN THE SPECTRAL PROFILES OF EACH COLONIC WALL LAYER. C) CONFOUND MATRIX DEMONSTRATING CLASSIFICATION ACCURACY OF THE MODEL................................................................................174

FIGURE 58: A) UNSUPERVISED (PCA) AND B) SUPERVISED (LDA) ANALYSIS DEMONSTRATING SEPARATION OF ADENOCARCINOMA AND NAM. C) CONFOUND MATRIC DEMONSTRATING CLASSIFICATION ACCURACY OF ENDOSCOPIC SNARE ANALYSIS OF COLORECTAL ADENOCARCINOMA (1) AND NORMAL ADJACENT MUCOSA (0) ...........................................................................................................176

FIGURE 59: LIPID SPECIES SIGNIFICANTLY ALTERED IN THE TRANSITION FROM NORMAL ADJACENT MUCOSA TO DYSPLASIA TO INVASIVE CANCER ....................................................................................................................184
Abbreviations

**AJCC** American Joint Committee on Cancer

**APC** adenomatous polyposis coli

**APR** abdominoperineal resection

**AUC** area under the curve

**BCSP** bowel cancer screening programme

**BRAF** B-Raf proto-oncogene, serine/threonine kinase.

**CEA** carcinoembryonic antigen

**CLE** confocal laser endomicroscopy

**CRM** circumferential resection margin

**DESI** desorption electrospray ionisation mass spectrometry

**EGFR** epidermal growth factor receptor

**EMR** endoscopic mucosal resection

**EMVI** extramural vascular invasion

**FOB** faecal occult blood

**FT-ICR** fourier-transform ion-cyclotron resonance

**HER2** Human Epidermal Growth Factor Receptor-2

**HR-MAS NMR** high-resolution magic angle spinning nuclear magnetic resonance

**IHC** immunohistochemistry

**IPA** isopropyl alcohol

**KRAS** V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog

**LDA** linear discriminant analysis

**MALDI** matrix assisted laser desorption ionization

**MMC** maximum margin criterion

**MS** mass spectrometry

**MSI** mass spectrometry imaging
**MSI** microsatellite instability

**m/z** mass-to-charge ratio

**NAM** normal adjacent mucosa

**NICE** National Institute for Health and Care Excellence

**NBI** narrow band imaging

**OPLS** orthogonal partial least squares

**PC** phosphocholine

**PE** phosphatidylethanolamine

**PI** phosphatidylinositol

**PS** phosphatidylserine

**PC** principal component

**PCA** principal component analysis

**RF** radiofrequency

**ROC** receiver operating characteristic

**SCFA** short chain fatty acid

**SIMS** secondary ion mass spectrometry

**TEM** transanal endoscopic microsurgery

**ToF** time of flight

**S-FU** 5-fluorouracil
1. Background

1.1 Principles of Oncological Surgery

Surgical oncological practice is founded on the principles of a clear tumour resection margin and surgical removal of associated vascular and lymphatic channels. This approach has remained unchanged for over 100 years, and it remains highly applicable to the modern management of cancer. Over 300,000 cancers are diagnosed in the UK each year and 1.8 million diagnostic, curative or palliative surgical procedures are performed as part of their treatment\(^1\). Surgical excision remains the gold standard of care for the majority of solid organ tumours and from an oncological perspective, a curative surgical resection is determined by the complete removal of tumour tissue with an associated border of histologically normal tissue. The principles of surgical oncology originated with John Hunter in the mid 18\(^{th}\) century. He first described many concepts of oncological surgery; including that cancer is a localised process, potentially amenable to surgical cure, and stressed the need for total removal of the cancer and potential areas of dissemination. These theories would not be realized in practice until this type of radical surgery became feasible through a better understanding of anatomy and pathology. Over 100 years later, in 1895, William Halsted, one of the founding physicians of John Hopkins Hospital and of the radical mastectomy, with which his name is synonymous, demonstrated that complete removal of all tumour tissue with draining lymphatics significantly reduced risk of tumour recurrence. Aligning with these principles, the Halstead radical mastectomy incorporated excision of all breast tissue, axillary lymph nodes and both pectoralis muscles. This approach, though oncologically sound, often resulted in severe disfigurement of the patient, weakened arm function and disabling lymphoedema.
Figure 1: a) Halsted radical mastectomy demonstrating removal of the breast, pectoral muscles, axillary lymphnodes and skin (image from Halsted. The Results of Operations for the Cure of Cancer of the Breast Performed at the Johns Hopkins Hospital from June, 1889, to January, 1894 Ann Surg. 1894 Nov; 20(5); 497-555) and b) pictorial representation of a modern wide local excision of breast cancer.

The role of surgery in the treatment of cancer has evolved significantly over the last century to include less invasive surgical approaches, but it is still based on Halsted’s fundamental oncological principles. The mid 20\textsuperscript{th} century saw major advances in cancer management outside the realms of surgery\textsuperscript{2} and, as mechanistic insights into tumour biology have been gained and next generation adjuvant treatments have been developed, the role of surgery has evolved from the sole modality for cure, to one aspect in a wealth of treatment strategies used in the multidisciplinary management of cancer. The increasing use of neoadjuvant therapy means that surgery is no longer always first line treatment, and surgical oncology routinely includes procedures for diagnosis and staging, palliation and prophylaxis. However, even in the contemporary era of multi-modal cancer therapy, it is rare for the care of a patient with a solid organ malignancy not to include a surgical component. As the role of surgery has changed, so therefore has the role of the surgeon; the modern surgical oncologist must have the necessary knowledge, skills and experience to manage individual patients through what is often a complex and long cancer patient journey. A fundamental component of this practice is based on accurate phenotyping of patient specific tumour biology, however other than a small number of largely genetic
screening tests, there are very few technologies that are able to relay this information in real time. For example, decisions regarding extent of surgical excision required (organ sparing vs. radical) may be based on limited information, as much of the detail required to determine who will benefit from either approach is provided during full histological examination of the resected specimen. However, if survival and quality of life after cancer surgery are to be improved there is a significant need to improve the quality of the primary surgical management of solid tumors. The most obvious area for improvement in operative oncology is the surgical resection margin. This is because the risk of local recurrence for solid malignancies is increased if negative margins are not achieved and, particularly with modern, less invasive approaches, there is a risk that microscopically, malignant cells may extend beyond the visible or palpable tumour border. Determining areas of this microscopic infiltration at the time of surgical resection is challenging; there is rarely a well-defined boundary that can be visualized intra-operatively and only when post-operative pathological examination is carried out can “close” or positive resection margins be identified. This may necessitate further surgical intervention to clear operative margins with resulting increased operative risk, increased use of resources, delay to adjuvant therapy, reduction in the quality of the patient experience, risk of complications and in certain cases a poorer cosmetic outcome.

There are currently no standard methods for providing accurate, real time information on the histological nature of tissue to the surgeon intra-operatively or during diagnostic investigations and procedures used to guide and plan therapy. Instead, the surgeon must make intra-operative decisions based upon visual and tactile information, or if in serious doubt about adequacy of resection, utilize the intra-operative frozen section technique in an attempt to identify residual tumour tissue. True adequacy of resection is determined on postoperative histological examination after the fact. Intra-operative tools that could identify positive resection margins and involved lymphatics could assist the surgeon in ensuring adequate dissection of tumour margins at the time of the procedure to avoid leaving residual disease. They could also provide information to determine whether a minimally invasive organ sparing approach is indeed a feasible oncological option for each individual patient and
allow intra-operative decisions about the extent of resection required for an oncologically safe outcome.

1.2 –Omics, Personalized Healthcare and Stratified Medicine

Omics technologies adopt a holistic view of the molecules that make up a cell, tissue or organism. They are aimed at the non-targeted universal detection of genes (genomics), proteins (proteomics) or metabolites (metabolomics) within a biological sample\(^3\) and differ from traditional experiments in not being hypothesis driven, but acquiring and analyzing all generated data to define a hypothesis that can be further tested\(^4\) – a top down approach. The integration of omics techniques is often referred to as Systems Biology\(^5\). Personalised genomics medicine, made possible due to increased knowledge of the human genome following completion of the Human Genome Project\(^6\) was the first omics science to be developed and applied to clinical medicine. DNA microarrays measure differences in DNA sequences between individuals and allow analysis of the expression of thousands of genes simultaneously. Knowledge of individual patients’ genotype can be applied to provide information on disease susceptibility and effectiveness of medical intervention. In 2006 the Genomics and Personalized Medicine Act\(^7\) was created, this aimed to increase the usefulness and application of genomic knowledge to allow physicians to use information from an individual’s genome and its components to influence the decision-making process regarding patient care. In 2008 the Council of Advisors on Science and Technology (PCAST) provided the first comprehensive definition of personalised medicine

“Personalised Medicine refers to the tailoring of medical treatment to the individual characteristics of each patient. It does not literally mean the creation of drugs or medical devices that are unique to a patient, but rather the ability to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment. Preventive or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not”\(^8\).
Subsequently, Genome Wide Association Studies (GWAS) were established to investigate common genetic variants associated with individual disease processes and a new genetic mapping consortium project was begun aiming to sequence, in full, the genome of 1000 individuals to identify the most common human genetic variants and to make the resulting data available to researchers for use in GWAS. Clinically significant molecular targets discovered through genomics studies include ER and HER2 receptor expression in breast cancer and the recognition of the KRAS mutation in colorectal cancer\(^9\). The ability to selectively modulate these targets and select appropriate therapy based on expected response has dramatically influenced personalized management of these diseases\(^10,11\).

Whilst remarkable progress has been made in genomic sequencing, translation into clinic practice has been hindered by lack of a practical model through which information acquired through genomics studies can readily be applied to clinical care\(^12\). Although GWAS have identified many novel genetic variants associated with pathogenesis of malignancy\(^13\), they do not always explain the association between genetic susceptibility and incidence of disease. Clinical outcomes are influenced by a large number of genomic, environmental, microbial, and clinical variables and genomics platforms alone are not able to account for these numerous environmental influences.

The molecular code within the genome influences cellular function through the creation of proteins. The complete set of proteins within a cell is referred to as the proteome, first described in 1994. Proteomics aims to characterise information flow within the cell and the organism through protein pathways and networks\(^14\). The proteome is dynamic and changes in response to environmental stimuli and is therefore, in contrast to the genome, a reflection of both genetics and the environment\(^3\). The Human Proteome Project, organised by the Human Proteome Organisation aims to map the entire human proteome to aid in the development of proteomics based diagnostic, prognostic and therapeutic medical applications.

The metabolome refers to the complete set of low molecular weight compounds in a biological sample. Metabolomics, the youngest of the omics approaches, has a number of proposed advantages over
other omics technologies. It is the final down stream product of gene transcription, and includes all cellular interactions including that of symbiotic pathways, and is therefore reflective of continuously changing metabolic and cellular signaling pathways. It is also closest to the phenotype of the biological system studied, being responsive to both host and environmental factors. The term metabolomics was first used in 1998 and the Human Metabolome Database, an online database of human metabolites was released in 2007 to facilitate metabolomics research.

The added value of genomics approaches in the medical management of cancer is made clear by the advantages that knowledge of genetic mutations such as BRCA, APC, RET, HER2 and KRAS bring to the stratification and medical management of cancer. Gene expression profiling assays such the Oncotype DX® in breast cancer profile show promising prognostic value and are beneficial as predictors of the value of adjuvant therapy for individual patients with cancer. Similar genotyping studies predicting response to adjuvant therapy have also been proposed as adjuncts in the decision making around timing and necessity of resectional surgery in e.g. rectal and oesophageal cancer where a complete clinical or pathological response may be seen. However, few other omics techniques are currently deployed in the surgical setting which impact surgical decision making. One genomics example, which has translated into the clinical domain, is the intraoperative one-step nucleic acid amplification (OSNA) assay for the identification of positive sentinel lymphnodes during breast cancer resections. The OSNA test is a molecular detection procedure that rapidly identifies lymph node metastases in breast cancer by detection and amplification of cytokeratin 19 (CK19) mRNA. This test is carried out intra-operatively, takes 40 minutes, using the whole lymph node and has been shown to detect more sentinel lymph node metastases, particularly micrometastases, than frozen section. No proteomics or metabolomics based technologies are as yet used routinely in the operating theatre environment.
1.3 Metabonomics

Metabonomics and metabolomics are terms that are often used interchangeably. Whilst metabolomics aims to analyse, characterize and quantify small molecules in a biological sample, metabonomics seeks to measure the global dynamic shifts in small molecules within a multicellular system in response to genetic or environmental change\(^{21}\). Metabonomics uses computational analysis of spectral metabolite data to investigate system wide time specific changes providing an overview of the metabolic state of an individual with all of its genetic and environmental influences at a given point in time – Nicholson et al. 1999. This approach measures the metabolic endpoints of whole system activity, which are influenced by genetic and environmental factors including parasitic and extended genome (such as gut microbial) interactions\(^{3}\). These gene-environment interactions can lead to multiple outcomes, and the metabolic phenotype of an individual represents the outcome of this interaction as measured by shifts in the profile of small molecules within a biological sample. Global metabolic profiling does not separate endogenous metabolic processes (those coded in the genome and intrinsic to cellular function) and response to environmental stimulus but rather the final outcome of the interaction between the two\(^{4}\).

Metabolic profiling uses high throughput spectral analytical platforms such as \(^1\)H Nuclear Magnetic Resonance spectroscopy (\(^1\)HNMNR), Magic Angle Spinning NMR and Mass Spectrometry, in conjunction with multivariate computational statistical analysis, to provide information on the phenotypic outcome of complex shifts in the metabolic profiles the biological samples analysed. Platforms such as \(^1\)H NMR spectroscopy or mass spectrometry rapidly measure large arrays of metabolites providing high density, complex data sets rich in molecular information which are robust and high throughput.

NMR spectroscopy identifies compounds by applying radio-frequency energy to samples within a magnetic field. The nuclei of certain elements such as hydrogen (\(^1\)H) align to the magnetic field, and changes in this alignment are detectable as a release of energy. Rather than the image generated by MRI, NMR spectroscopy produces spectra of peaks, corresponding to chemical groups within compounds. Until recently, NMR analysis was limited to fluid samples however the development of
magic angle spinning $^1$H NMR spectroscopy has broadened the capacity to include intact tissue samples, which can be analysed in a rapid and non-destructive manner.

The role of mass spectrometry in endogenous metabolite research is evolving with the development of new technology. MS is well established as a tool for quantifying small molecules and it is a valuable tool for the identification of metabolites and biomarkers. Mass spectrometry has traditionally been used for the analysis of biofluid samples however more recently techniques have been developed for the analysis of tissue samples. Traditionally the need for complex sample extraction and chromatographic separation prior to MS analysis has precluded its translation to the clinical environment, however, the development of direct injection and ambient ionisation methods using highly simplified procedures which consume very little sample could lead to MS becoming a tool for clinical biofluid and tissue analysis in the near future. Mass spectrometry imaging methods (MSI) such as MALDI and DESI create images of tissue samples based on mass-to-charge ratios, generating mass spectra correlated to known locations within the tissue, thus forming metabolite-specific images demonstrating the topographical distribution of metabolites across tissue sections. This information can be used to determine the metabolic differences between tissue types within a single sample (such as tumour tissue, tumour supporting stroma and normal epithelial tissue), drug distribution throughout a tissue specimen and for microbial identification.

MALDI MSI has recently been introduced to the clinical environment as a tool for histological and microbiological tissue classification. MALDI MSI uses a matrix to aid laser desorption ionisation of tissue specific metabolites from prepared tissue sections. This generates characteristic mass spectral profiles specific to each microorganism, which can be used for microbial identification. The fact that unique MALDI spectral profiles can be obtained from whole bacterial cells without pretreatment was first reported in 1996 and its use in clinical microbiology laboratories is increasing. In practice, MALDI analysis of microorganisms involves the mixture of a microbial sample with a matrix on a conductive plate. After introduction into the mass spectrometer it is bombarded with laser pulses, the matrix absorbs energy from the laser, which leads to desorption of sample specific analytes which are then analysed in the gas phase. The spectra generated can then be compared to commercially available
spectral libraries for species identification. The major advantage over traditional methods of identification such as microbial growth, colonization and gram staining is a significantly reduced time to result (hours), in addition it allows identification of a large variety of bacteria with few phenotypic traits which previously relied upon 16S rRNA gene sequencing.

DESI MSI is a further MS imaging method which provides high resolution, spatially resolved metabolite based images of tissue sections under ambient conditions without the need for a matrix or any tissue preparation. DESI utilises an electrospray to promote the formation of tissue specific ionised gas phase particles which are suitable for analysis by the spectrometer. DESI was first described as an imaging method in 2004, and currently remains a research laboratory tool, used to investigate small molecule changes in tissue samples, associated with disease processes such as malignancy but has potential as a future tool for histological tissue diagnosis in the clinical domain. The instrumentation and analysis methods used in mass spectrometry analysis are described in more detail in Chapter 2.

Metabonomics analysis strategies are useful tools in biomarker discovery as vast multitudes of molecules are investigated simultaneously. Biomarkers can be classified as diagnostic, prognostic (predict survival independent of treatment effects) or predictive (predict benefit from a particular therapy). Biomarker discovery has traditionally focused on the identification of proteins that are dysregulated as a consequence of the disease and are shed into, and are detectable in biofluids. This approach relies upon each candidate marker being identified among thousands of proteins and is thus an arduous process. Advances in metabonomic technologies show promise in the discovery of new biomarkers or, more likely, panels of biomarkers, which may be used not only in diagnosis of malignancy but also in prognosis and prediction of response to treatment, identification of new therapeutic targets and improved understanding of the molecular alterations in malignancy.

1.4 Computational Systems Medicine and Bioinformatics

The significant complexity and size of data produced by omics based technologies require complex multivariate modeling techniques for data interpretation, analysis and visualization. Chemometric techniques used in untargeted metabonomics analysis fall into two broad categories; unsupervised and
supervised. Unsupervised techniques search for clustering without prior knowledge of the data, and include linear techniques such as principal component analysis (PCA), and non-linear techniques such as hierarchical cluster analysis (HCA) and self-organising maps (SOMs). These techniques use simplification or dimensionality reduction to identify the inherent variation within the dataset using a reduced number of parameters with minimal loss of information. Supervised techniques use a training set of data, whose classification is already known, to predict the class of a separate validation set. Techniques include linear discriminant analysis (LDA), Maximum Margin Criterion analysis (MMC), partial least-squares (PLS) regression and variants such as partial least-squares and orthogonal partial least squares discriminant analysis (PLS-DA and OPLS-DA). Example PCA and LDA plots are shown in figure 2. Output models use 3 axes: length, width and height to represent a 3 dimensional dataspace. When datapoints are plotted, the geometric centre of the dataspace is calculated and new coordinate axes are created, orthogonal to each other, pointing in the direction of the largest variance. This produces fewer variables as each datapoint can be described by its position along a component rather than by length, width and height, but the relative geometric position of each data point remains unchanged and no information is lost. Components are numerically ordered by the amount of variance of the datapoints along them and therefore Component 1 represents the direction of the largest variance, component 2 the second largest and so on. Whilst principal components are the axes which represent the most variance within the dataset, linear discriminants are the axes which most strongly differentiate the individual classes; therefore LDA maximizes separation between each class within the dataset. PCA and LDA can be used independently for multi-class classification tasks but are often used in combination e.g. PCA for dimensionality reduction followed by LDA for class separation. This combination of unsupervised analysis by PCA followed by supervised analysis by LDA is used frequently throughout this body of work for analysis of spectral data.
Figure 2: a) Principal components analysis (unsupervised) of a dataset with 5 individual classes. Maximal variance within the data can be seen along PC1. b) Linear discriminant analysis (supervised) identifies the axes responsible for the greatest differentiation between classes and has maximised separation between the groups.

One of the most difficult tasks in a top down systems biology approach lies in the identification of biologically meaningful metabolites and subsequent correlation with a disease pathogenesis or physiological/metabolic pathways. For this to be successful, metabolite identity must be assigned to individual features, found to be significant in the process of untargeted analysis. Improvements in methods for visualisation of results and of filtering features have facilitated characterization and structural identification, however, this remains a time consuming, complex part of the process. Tandem MS data is first acquired for the feature of interest, and this is then compared to metabolite databases such as METLIN\textsuperscript{26} for definitive identification and validation. Although some new online processes are available to automate and speed up the process, most metabolite identification still requires manual searching. Once individual metabolites have been identified, pathway and network tools such as KEGG\textsuperscript{27} and the Edinburgh Human Metabolic Network\textsuperscript{28} can be used to elucidate the role of identified metabolites in biochemical networks and pathways.
1.5 Application of Metabolic Phenotyping in the Clinical Environment

An individual’s metabolic phenotype is the product of genetic and environmental factors at a specific moment in time under a specific set of conditions. Most disease processes are subject to environmental influences and these external influences also affect aspects of clinical care, such as drug metabolism. Metabolic phenotypes are therefore also a product of, and responsive to, medical and surgical intervention. Introduction of metabolic profiling tools into the clinical or surgical environment could enhance clinical decision-making, based on this highly personalised, detailed molecular information and provide therapies that are tailored to the biological state of the individual patient at each stage of their clinical treatment. Integration of top down systems biology tools with conventional clinical diagnostic tools and patient information could be used to optimise risk stratification, healthcare planning and personalized therapeutics.

A significant goal of next-generation personalised technologies relates to the improved preoperative prediction of postoperative outcome, to permit better risk stratification and patient selection, improve pre-operative physiological optimisation and allow appropriate targeting of supportive care. Similarly, it has been proposed that longitudinal patient profiling based upon biofluid analysis could be used in the early identification of post operative complications. Oncological surgery could also benefit significantly from improved tissue based diagnostics, more accurate staging of disease and prediction of tumour characteristics, all of which influence decision making in the surgical management of cancer. Tissue based profiling techniques such as mass spectrometry imaging (MSI) have the potential to provide high-throughput, low cost per sample, rapid tissue profiling and could provide a more efficient, accurate alternative to intra-operative frozen section for the analysis of tissue sections. In addition, metabolic profiling of tissue samples can provide additional information on molecular alterations in cancer tissue as well as prediction of prognostic characteristics such as aggressive clinical features and lymphatic involvement. Profiling tools have also identified the presence of cancer associated metabolic change in histologically normal, tumour adjacent tissue supporting the theory of field cancerisation, first described in squamous cell oral cancers, that early genetic events that lead to the morphologically apparent malignant transformation of cells also leads to a population of peritumoural cells with early
genetic (and biochemical) changes towards malignant transformation but without apparent histological changes\textsuperscript{30}. A better understanding of the inherent molecular heterogeneity in cancer is required for improved operative outcomes and judicious use of targeted adjuvant therapies. There are few molecular techniques available to the surgeon in the clinical setting to provide diagnostic or prognostic information. Over the last decades, advances in adjuvant systemic therapy have substantially improved the management of patients with a poor prognosis at the time of diagnosis. However, despite appropriate surgical and systemic therapy, current clinical and pathological parameters are unable to identify, among this population, those patients for whom the outcome of current standard treatment will be curative and those who will ultimately relapse. Such a failure in prediction may be due to a relative inability to fully address the molecular heterogeneity inherent to the process of malignant transformation of tissue and this represents an obstacle to a the more personalized management of cancer. Systems medicine platforms could be employed in the surgical and oncological management of cancer from providing insight into the aetiopathogenesis of disease and tumour characteristics to prediction of outcome, risk of recurrence and response to therapeutic agents, as well as the identification of new therapeutic targets. Novel, near real time tissue diagnostic strategies also have potential to improve patient outcome by reducing the incidence of positive resection margins during cancer resections and providing information on the biochemical nature of individual tumours \textit{in vivo} to augment the surgical decision making process.

New systems-level technologies are being developed to be robust, cost effective, rapid and translatable into the clinical environment. However, in order for global metabolic profiling techniques to fully translate into clinical practice, sophisticated computational interfaces and data visualization platforms are required, as well as better designed clinical and pathology databases into which this relatively complex data can integrate. In reaction to the growth of metabolic profiling as a potential clinical tool, novel bioinformatics platforms with more user-friendly interface tools are also being developed to handle the significant volumes of data produced\textsuperscript{10}.

This thesis presents a novel method for \textit{in vivo} metabolic profiling of cancer tissue using mass spectrometry analysis of the electrosurgical smoke plume during tissue dissection. This method, known
as Rapid Evaporative Ionisation Mass Spectrometry or the iKnife, is proposed as a future tool to provide near real time, intra-operative tissue diagnostics during surgical resection and improved phenotyping of cancer and stratification of disease based upon multivariate analysis of tumour lipid signatures. Breast and Colorectal cancer are the 2nd and 3rd most common cause of cancer deaths in the UK respectively and both pathologies have an unmet need for near real time tumour phenotyping in the operative and clinical environment. This thesis focuses on the application of REIMS to both colorectal and breast cancer.

1.6 Colorectal Cancer

Colorectal cancer is the second most common cancer in Europe, with around 447,000 new cases diagnosed in 2012 (13% of the total number of cancers). The incidence of colorectal cancer increases with age, with greater than 95% occurring in people aged over 50 years and rates of diagnosis have increased by 6% in the last 10 years. Slightly more men than women are affected. More than 90% of colorectal cancers are histologically adenocarcinoma, a malignant tumour formed within the glandular structures of epithelial tissue, but rare tumour types include carcinoid, lymphoma and neuroendocrine tumours. 27% of colorectal cancers occur in the rectum, 7% at the rectosigmoid junction, 20% in the sigmoid, 14% in the caecum, 21% in the remainder of the colon, 2% in the anal canal, 1% in the appendix, (in 8% location is unspecified). The majority of colorectal adenocarcinomas arise from the adenoma-carcinoma sequence, the stepwise progression from normal to dysplastic epithelium to carcinoma associated with multiple selected genetic alterations. Certain single-gene disorders lead to an increased risk of colorectal cancer and account for about 10% to 15% of cases. Hereditary non-polyposis colorectal cancer (HNPCC) is the most common genetic disorder and accounts for approximately 3-5% of colorectal cancers. It results from defects in mismatch repair genes involving MSH2, MLH1, PMS1, PMS2, or MSH6. Other hereditary conditions include polyposis disorders, such as familial adenomatous polyposis, which occurs due to mutation of the APC tumour suppressor gene located on Chromosome 5q21 and leads to the formation of thousands of dysplastic polyps within the
colon with the associated risk of malignant transformation, and hamartomatous disorders such as Peutz-Jeghers syndrome.

Colorectal cancer is managed surgically with (neo)adjuvant chemotherapy/radiotherapy as required depending on tumour stage. Exceptions are locally advanced disease for which surgical resection will not provide adequate tumour clearance, although these tumours may still undergo debulking or resectional surgery to avoid symptomatic intestinal obstruction. Colorectal cancer is staged using the TNM (Tumour Nodes Metastases) 7TH edition staging system (See Appendix 1). TNM groupings are then combined and each individual is given a cancer stage based on the combination of each TNM component (See Appendix 2). Tumour stage determines prognosis, management, and who will benefit from systemic chemotherapy or radiotherapy.

1.6.1 Diagnosis and Early Detection

The clinical and economic burden of colorectal cancer on the UK NHS is increasing. Lower gastrointestinal endoscopy and biopsy remain the principle method for the diagnosis of intraluminal colorectal lesions and in the Bowel Cancer Screening Programme (BCSP) era it is estimated that up to 650,000 lower gastro-intestinal endoscopies (colonoscopy/flexible sigmoidoscopy) are performed annually in the UK, and there is a significant year on year increase. In the UK, faecal occult blood (FOB) testing is employed as the as the initial screening test for colorectal cancer. Of every 1,000 people who undertake FOB testing, around 20 (2%) will have a positive result, 16 of those will go on to have a colonoscopy 8 of which will be normal, 6 will have polyps, 2 will have cancer. A significant factor in the prevention of CRC is the endoscopic removal of precursor adenomatous polyps. Colonic polyps may be dysplastic, have potential for malignant transformation, or hyperplastic, truly benign growths with no malignant potential. Small or pedunculated polyps are removed by snare polypectomy, larger lesions can be removed by endoscopic mucosal resection. Any polypoid lesion seen at endoscopy must be excised if possible (biopsied if not) and sent for histological examination. If found to be dysplastic rather than hyperplastic in nature the individual must enter into a programme of routine colonoscopic surveillance, as after removal 30-35% of these patients will have further adenomas detected at 3-4
years. The introduction of screening has led to a substantial increase in the detection of benign colonic polyps; in a 3 year analysis of the BCSP 49,054 polyps were identified, of which a very small number (1.9%) were malignant (by size 0–9 mm 0.3%, 10–19 mm 4.4%, 20–29 mm 8%)36. This leads to a significant burden on histopathology and outpatient services, as each endoscopically biopsied sample must be formally processed and examined and the results relayed to the patient in an outpatient appointment typically 2 weeks after endoscopy. Due to the risk of malignant transformation, dysplastic polyps must be removed in their entirety during endoscopic examination and if this is not possible, surgical resection may be performed for large lesions, which carry increasing malignant potential. Conversely, hyperplastic polyps pose no risk of malignant transformation and as such can be left in situ. However, unenhanced endoscopic differentiation of hyperplastic and dysplastic polyps is unreliable and as such, all lesions seen at endoscopy are currently removed or biopsied and sent for histological examination.

Introduction of the BCSP has led to the detection of increasing numbers of early stage cancers. There has been a recent move towards endoscopic local management of these early malignant lesions, where possible, as a less invasive, organ sparing alternative to radical surgical resection. Endoscopic mucosal/submucosal resection (EMR) and transanal endoscopic microsurgical (TEM) techniques provide a method of local management of large dysplastic and early stage malignant lesions where disease is confined to the submucosal layer with an absence of lymphatic involvement. If there is any doubt about the TNM stage of the tumour or endoscopic management carries a risk of incomplete resection then full surgical resection is carried out. The incomplete resection rate for endoscopic resection of neoplastic polyps has been reported at 10% and this rate increases with increasing size of the lesion, being as high as 47% in sessile serrated adenomas between 10 and 20mm in size37. A recent pooled multi-cohort analysis concluded that up to 19% of colorectal malignancy arising after a previous colonoscopy may be due to the incomplete resection of an earlier non-invasive lesion38. In addition, increasing rates of advanced interventional endoscopy are associated with higher rates of colonic perforation, a significant complication, when compared to less invasive procedures39.
1.6.2 Surgical Management

Radical surgical resection with associated lymphadenectomy remains the cornerstone in the management of malignant colorectal disease. Local recurrence rates in rectal cancer have fallen dramatically over the past 20 years, with the recognition of the importance of CRM (circumferential resection margin) involvement and the subsequent global introduction of total mesorectal excision and neoadjuvant therapy\(^\text{40}\). Modern total mesorectal excision for rectal cancer involves excision of not only the rectum, but the entire mesorectal compartment including the rectum, surrounding mesorectal fat, perirectal lymphnodes and the mesorectal fascia. The plane of dissection is the circumferential resection margin, which must be preserved to ensure radical resection of all of the above structures. Failure to achieve this, or tumour involvement in this margin, results in higher risk of having a positive resection margin. In colorectal cancer, as with most solid organ malignancies, a positive resection margin leads to significantly poorer clinical outcomes and clear resection margin remains an important prognostic indicator for disease-free survival and is a qualitative marker of the surgical treatment of rectal cancer\(^\text{41}\). Cancer resection margins are described as R0 (no residual tumour at the resection margin), R1 (microscopic residual tumour at the resection margin) and R2 (macroscopic residual tumour at the resection margin). Still, malignant CRM involvement following rectal resection has been reported as being as high as 10.4% and 23.6% in those undergoing elective or emergency resections respectively\(^\text{42}\). A rectal resection can be considered as R1 based on 3 criteria: distal margin involvement, a circumferential resection margin less than or equal to 1 mm, or a circumferential resection margin with involved lymph nodes within 1 mm. Although an R0 resection also remains the optimum outcome at surgery for locally advanced and recurrent rectal malignancy, this is only achieved in approximately 57 to 69% of cases, while 20 to 25% are R1 and 11 to 16 % are classed as R2\(^\text{41}\). Surgery for locally advanced (requiring resection beyond the TME plane to achieve R0 resection) or recurrent rectal cancer requires extensive dissection within confines of the pelvis and negative resection margins are limited, in many cases, by anatomical boundaries. However, in potentially resectable locally advanced and recurrent rectal cancer, R0 resection margin rates are reported at 37% to 57%\(^\text{43}\), in part influenced by distorted tissue planes and long course neoadjuvant radiotherapy. More recently the importance of
prevention and management of distant metastases, as well as local recurrence, in rectal cancer has been recognized. Despite the improvement in local control associated with neoadjuvant therapy and standardized surgical techniques in locally advanced but resectable colorectal cancer, this has not been associated with increasing survival\(^{44,45,46}\). More recently the importance of prevention and management of distant metastases, as well as local recurrence, in rectal cancer has been recognized. Recent data suggests that, although local recurrence remains a clinically significant problem with significant morbidity and is difficult to salvage, metastasis may be the predominant problem with the incidence of distant metastases being 24-28%\(^{44}\). Data from a retrospective case matched study\(^{47}\) comparing outcomes between R0 and R1 resections matched for clinical characteristics, found that, whilst no significant increase in local recurrence was identified following a positive CRM, risk of distant metastases was significantly increased and in these cases, risk of metastases was significantly influenced by the provision of adjuvant chemotherapy. Notably, a positive CRM was also associated with certain aggressive tumour characteristics, particularly T stage.

With advances in modern surgical and endoscopic techniques and improved efficacy of adjunct therapies, more minimally invasive, organ preserving approaches are becoming acceptable management strategies for malignancies with specific clinical characteristics. In early rectal cancer, surgical treatment strategies depend heavily on depth of tumour invasion and nodal involvement. T1/2 N0 tumours can be successfully managed with organ sparing local resection such as transanal endoscopic microsurgery or endoscopic mucosal resection. As mentioned, introduction of the BCSP has led to the detection of increasingly early cancer with pilot studies reporting that 49-52% of screen detected rectal cancers are T1/2N0/M0/stage 1 at diagnosis\(^{48}\). However, despite the introduction and availability of minimally invasive local procedures, total mesorectal excision (TME) is commonly carried out for early stage tumours due to concern over undertreating those with nodal involvement, who will not be cured with local resection alone. Even during minimally invasive procedures such as transanal endoscopic microsurgery and endoscopic mucosal resection, principles of complete resection remain fundamentally important to reduce the risk of local recurrence and distant metastases. Formal lymphadenectomy does not form part of such minimally invasive procedures, therefore those with
disease more locally advanced than T2 or with lymphnode involvement are not suitable candidates. Whilst early cancer diagnosis has a positive impact on outcome, there remain several unknowns in the optimal endoscopic and surgical management of these more frequently encountered early tumours. Adequate case selection requires accurate tumour assessment for risk of nodal involvement and likelihood of achieving a negative margin in the pre- or even intra-operative phase to guide surgical management. Currently most information which guides clinical judgment on the need for adjuvant therapy or further surgical intervention is based upon histological assessment of the resected specimen after the fact. Successful pre-operative differentiation, in the form of biomarkers of disease stage and nodal involvement are required to aid identification of those who can be treated successfully with an organ sparing approach prior to the decision making process. Additionally, post-operative histopathological examination of locally excised lesions lacks precision for the identification of those for whom conversion to TME would have been optimal49. In some cases, neoadjuvant radiotherapy prior to local resection could reduce the risk of recurrence, and downstaging followed by local resection may be an alternative to radical surgery, but identifying those individuals who are suitable in the pre-treatment phase is challenging. Outcomes of the addition of adjuvant radiotherapy for locally resected tumours with less favourable histological characteristics are not satisfactory49,50. Despite this, most emerging tests are designed to personalise the post surgical treatment of colorectal cancer. Improved methods for in vivo, oncologically and histologically accurate tissue characterisation could be used not only in the surgical or endoscopic assessment of resection margins but in pre-operative diagnostics and assessment of dysplastic and malignant lesions to determine optimal treatment regimes prior to histological assessment of the resected specimen.

1.6.3 Prognosis, Staging and (neo)Adjuvant Therapy

Those diagnosed with high risk stage II or stage III colon cancer will be offered adjuvant chemotherapy. Occasionally neoadjuvant therapy may be used to downstage tumours prior to surgery. In moderate risk rectal cancer, neoadjuvant chemoradiotherapy is more commonly used to downstage the tumour prior to surgery, followed by adjuvant chemotherapy if required. Palliative chemotherapy may be
offered to those with disseminated or unresectable disease to relieve symptoms and prolong survival. Pre-operative T stage, as well as nodal and CRM involvement form the bases of assessment for neoadjuvant therapy. Guidelines and practice vary globally\textsuperscript{51} but in the UK, patients with T3 or T4 disease and those with N1 or N2 disease are considered for neoadjuvant therapy (NICE clinical guideline 131). Modern local staging of rectal cancer has evolved significantly with the advent of MRI\textsuperscript{52}. MRI is the current gold standard imaging modality used for pre-operative assessment of rectal tumour characteristics including T stage, lymph node involvement and CRM involvement. The presence of extramural venous invasion (EMVI) and CRM involvement, as seen on MRI are additional factors used to determine risk and certain MRI based imaging biomarkers have been proposed as independent predictors of outcome or response to neoadjuvant therapy in rectal cancer, as a strategy for further stratifying these patients prior to surgery\textsuperscript{53,54}. However, a recent meta-analysis demonstrated the specificity of MRI for predicting T stage, lymph node involvement and CRM involvement to be relatively low at 75% and 71% and 94% when compared to post-operative histopathology with a sensitivity of 87%, 77% and 94\textsuperscript{55}.

1.6.4 Biomarkers

The first recognized test for a common cancer, or biomarker, was carcinoembryonic antigen (CEA), reported in 1965 as a serum diagnostic marker of colorectal cancer by Dr Joseph Gold\textsuperscript{11}. Since that time, multiple serum detectable cancer biomarkers have been introduced into clinical practice including Prostate Specific Antigen (PSA) CA 125 and calcitonin. These markers have proved less useful in the diagnosis of, or screening for, cancer due to poor specificity but are routinely used in monitoring for disease recurrence and also response to treatment. Multiple histopathological markers are used for treatment selection for (neo)adjuvant therapy. These include tumour grade, stage, nodal involvement, lymphovascular and perineural invasion\textsuperscript{56} and microsatellite instability due to loss of DNA mismatch repair activity. However, there are no biomarkers for guiding surgical strategies designed to minimise the morbidity of surgery. This is particularly pertinent in the rectum where because of key anatomical and biological differences, adenocarcinoma is staged and managed differently. New cancer biomarkers
that not only detect cancer but predict outcome and influence treatment choice are required to assist the surgeon at the pre- or intra-operative decision making stage, particularly in determining who will benefit from radical surgical resection or systemic therapy. Current biomarkers proposed for inclusion in the management of colorectal cancer include tissue based markers such as EGFR, KRAS, BRAF, Ki 67, p53 and p21 and imaging based biomarkers such as extramural venous invasion (EMVI). The presence of EMVI, as seen on MRI, is both a prognostic indicator and a predictive marker of treatment response, as regression of more than 50% following neoadjuvant chemoradiotherapy is associated with an improved disease free survival compared to those with regression of less than 50%\textsuperscript{54}. EGFR overexpression, whilst associated with poor survival in other malignancies, is not universally confirmed as in independent predictor of prognosis in colorectal cancer as the impact of the EGFR signaling pathway on colorectal carcinogenesis in not fully understood\textsuperscript{57}.

KRAS mutations are a major predictor of resistance to anti-EGFR monoclonal antibody based chemotherapeutic agents. To improve outcomes and minimize unnecessary toxicity and cost, the National Comprehensive Cancer Network, the European Society for Medical Oncology and American Society of Clinical Oncology recommend the use of anti EGFR directed therapy only in those with wild-type KRAS status. However, not all KRAS mutations have the same impact on mediating EGFR resistance and other RAS mutations may have a similar effect\textsuperscript{58}. High Ki-67 expression (a proliferation marker of expression of Enhancer of zeste homologue 2 gene) is associated with an increased disease free survival in colon (not rectal) cancer patients who received adjuvant chemotherapy but is not significant in those who received surgery alone and thus could be a predictive marker of the effect of adjuvant chemotherapy in those with stage III colon cancer\textsuperscript{59}. In comparison to stage III colon cancer (lymph node involvement), where the addition of chemotherapy has conclusively been demonstrated to improve disease free survival, in those with stage II disease (transmural bowel wall involvement but negative lymph nodes), there is no definitive evidence as to who may benefit from adjuvant treatment. Nice guidelines (Clinical Guideline 131) state that adjuvant chemotherapy should be considered for patients with high-risk stage II colon cancer, but this cohort is poorly defined. Adjuvant chemotherapy targets micrometastatic disease which, whilst beneficial in the improvement of disease free and overall
survival in those with stage III disease, has an undefined role in the management of stage II due to wide differences in prognosis within this disease classification\textsuperscript{60,61,62}. Histopathological prognostic markers such as T stage (T4 vs. T3)\textsuperscript{63}, poorly differentiated/high grade cancers\textsuperscript{64}, low lymph node retrieval and lymphovascular invasion\textsuperscript{64} may signify higher risk, and stage II disease has now been reclassified into stage II/T3 and stage II/T4 due to significant differences in 5 year survival between T3 and T4 tumours within stage II disease\textsuperscript{63}. However, patients without any of these features of poor prognosis are classified as average risk, where the benefit of adjuvant therapy is unproven.

Molecular differences may exist between stage II and stage III or indeed between moderate and high risk stage II disease and better understanding of the biology of stage II cancer may lead to molecular markers which could improve risk stratification in this subgroup of patients or predictive markers of response to adjuvant therapy. Current potential molecular markers include high levels of microsatellite instability (MSI-H), which whilst associated with increased survival in this subgroup of patients, has not been validated as a predictive marker of benefit from chemotherapy\textsuperscript{65}. Thus the value of MSI-H as a tool in the management of patients with stage II disease who harbor any traditional pathological markers of poor prognosis is unclear. Loss of heterozygosity of chromosome 18 (18qLOH) has also been shown to be a poor prognostic indicator in patients with stage II disease\textsuperscript{66}. Management of moderate risk locally advanced rectal cancer, whether patients should undergo neoadjuvant therapy, and whether this should be with short course radiotherapy or chemoradiotherapy is also unclear. Pre-operative MRI assessment of established pathological risk factors including CRM involvement, lymph node involvement and extramural venous invasion\textsuperscript{67} allows stratification into low and high risk groups, with low risk patients suitable for surgical management alone. However further biomarkers (tissue and imaging) are required to predict the efficacy of neoadjuvant therapy to optimize management in those classified in the moderate-high risk group, particularly for cases of T3 rectal cancer with no further poor prognostic features (extramural spread <5 mm and without nodal or extramural venous involvement) for whom the survival benefit of neoadjuvant therapy remains unclear. A summary of current and potential future biomarkers for use in screening, diagnoses and management of colorectal cancer is provided in table 1.
<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Clinical Use</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEA (Carcinoembryonic Antigen)</td>
<td>Serum biomarker used clinically in surveillance for disease recurrence.</td>
<td>CEA is a glycoprotein involved in cell adhesion. It has poor specificity as a diagnostic or screening biomarker as elevated levels can be found in heavy smokers, multiple malignant (colorectal, pancreatic, gastric, lung and breast cancer) and benign conditions (diabetes, ulcerative colitis, pancreatitis and liver cirrhosis).</td>
</tr>
<tr>
<td>Mutation in one of the mismatch repair genes MLH1, MSH2, MSH6 or PSM2</td>
<td>Molecular biomarkers of Lynch Syndrome detected in resected tumour samples or in plasma.</td>
<td>Lynch syndrome is an autosomal dominant condition with variable penetrance which leads to a lifetime risk of CRC of 80% if not screened and treated.</td>
</tr>
<tr>
<td>KRAS and BRAF</td>
<td>Tissue molecular markers of predicted response to anti-EGFR chemotherapy.</td>
<td>KRAS encodes a g-protein, and BRAF encodes a protein kinase. Both are part of the ras/raf/mitogen-activated protein kinase intracellular pathway which stimulates the expression of proteins involved in cell division and survival. The presence of a KRAS mutation predicts a complete lack of response to anti-EGFR therapy. KRAS mutation is also an independent predictor of poor outcome regardless of tumour stage or intervention. KRAS mutation accounts for only 30–40% of patients who do not respond to anti-EGFR therapy. BRAF mutation, as an effector of KRAS, may also have a predictive role in the response to therapy with anti-EGFR in cancers in patients with wild-type KRAS.</td>
</tr>
<tr>
<td>Ki 67</td>
<td>Tissue based predictive marker of the benefit of adjuvant chemotherapy in stage III colon</td>
<td>Ki-67 is a monoclonal antibody that recognises an antigen present in the nuclei of cells in all phases of the cell cycle except G0. It is used as a marker of tumour proliferation and high expression is associated with an increased disease free survival.</td>
</tr>
</tbody>
</table>
cancer. survival after adjuvant chemotherapy in stage III colon cancer.

**p53**

Tissue based potential prognostic marker of 5 year survival and predictor of response to adjuvant chemotherapy.

The p53 tumor suppressor gene is the most frequently mutated gene in human cancer, seen in up to 50% of tumours of diverse histological types. p53 acts as a key regulator of cell growth control and plays a central role in the induction of genes that are important in cell cycle arrest and apoptosis following DNA damage\(^71\). There is no general conclusion on the prognostic or predictive role of P53 in colorectal cancer yet but p53 mutation has been associated with significantly worse 5-year survival\(^72\) and is also associated with distal and rectal cancers, lymphatic invasion and poorer survival following adjuvant chemotherapy\(^73\).

**p21**

Tissue based potential predictive marker of response to chemotherapy

p21, a cell cycle inhibitor, is transcriptionally regulated by p53 and mediates p53 dependent growth arrest. Additionally, p21 acts as an effector of multiple tumour suppressor pathways independent of the p53 tumour suppressor pathway\(^74\). The prognostic impact of p21 expression in colorectal carcinomas is still under debate. There is some evidence that high expression is associated with poor response to chemotherapy in rectal cancer\(^75\).

**EMVI (Extramural venous invasion)**

MRI and tissue based prognostic marker and predictive marker of response to neoadjuvant therapy.

In use

EMVI is defined by evidence of tumour cells in the vasculature outside the muscularis propria. It is a prognostic indicator associated with decreased disease free survival. Evidence of EMVI regression on MRI is also a predictive marker of response to neoadjuvant therapy.

**High levels of microsatellite instability (MSI-H)**

Tissue based prognostic biomarker.

There is consensus that a high levels of microsatellite instability are associated with improved prognosis even though MSI tends to be associated with poorly differentiated tumours\(^76\).

**Loss of heterozygosity of chromosome 18 (18qLOH)**

Tissue based prognostic biomarker.

Chromosome 18 contains several important genes involved in carcinogenesis and 18qLOH is associated with chromosomal
instability. Clinically, in colorectal cancer, 18qLOH is associated with tumours on the left-side of the colon and poorly-differentiated tumours. It is also an independent poor prognostic indicator in patients with stage II disease.66

<table>
<thead>
<tr>
<th>DYPD(dihydropyrimidine dehydrogenase) mutations</th>
<th>Serum based marker of increased risk of 5-FU toxicity</th>
<th>DYPD is a hepatic enzyme involved in the catabolism of uracil and thymine and is the initial rate-limiting enzyme in the metabolism of 5-FU in the liver, therefore deficiency is associated with increased toxicity secondary to impaired metabolism.77 Current predictive adequacy is not sufficiently accurate for screening of all patients prior to therapy78</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypermethylation of the Septin 9 gene</td>
<td>Plasma screening test</td>
<td>Septin 9, located on chromosome 17, has been shown to be hypermethylated in specimens of colorectal cancer tissue and can be detected in plasma in circulating DNA fractions of colorectal cancer cells. As a screening tool the test has a sensitivity of 72% and specificity of 90% for the detection of CRC.79 Multicentre clinical trial (PRESEPT trial) is currently underway.</td>
</tr>
<tr>
<td>Panels of biomarkers that include p53, KRAS, APC, BAT-26 and long DNA</td>
<td>Stool based DNA screening test for colorectal cancer and dysplastic adenomas</td>
<td>Panels of stool based markers have been proposed as better screening tools than individual markers. Sensitivity for the detection of cancer and adenomas of 52–91% and 15–82% respectively, and specificities of 93–100%68 These tests are expensive and currently the cost for population screening would be too high80</td>
</tr>
<tr>
<td>M2-PK (M2 Pyruvate Kinase) levels</td>
<td>Stool DNA screening test</td>
<td>Tumour M2-PK, an isoform of the glycolytic enzyme pyruvate kinase, is found in proliferating tissues with high nucleic acid synthesis such as tumour cells. Detection in stool samples as a screening test for colorectal cancer had a sensitivity and specificity of 72–91% and 78–79%68</td>
</tr>
<tr>
<td>Circulating free tumour mRNA</td>
<td>Plasma biomarker for surveillance of disease recurrence</td>
<td>Detection of free cancer cell mRNA in the plasma of colorectal cancer patients is independent predictor of postoperative recurrence.81 Currently this approach requires mutation profiling of each patient’s primary tumour</td>
</tr>
</tbody>
</table>
prior to detection, and therefore is costly and time-consuming\textsuperscript{68}.

### 1.7 Breast Cancer

Breast cancer is the most common cancer in European women accounting for 28\% of all cancers in the Western World. In 2011, 50 285 women in the United Kingdom alone were newly diagnosed with invasive breast cancer, many of whom underwent surgery with curative intent\textsuperscript{82}. Female breast cancer incidence rates in the UK have increased by 72\% since the mid 1970’s and 8 out of 10 cases are diagnosed in women over 50 years of age\textsuperscript{82}. Over 2 million women are screened each year as part of the national screening programme and this pathway alone detects over 17000 cancers 80\% of which are invasive, 19\% non invasive and 1\% micro-invasive at the time of diagnosis\textsuperscript{83}. More than 80\% of breast cancers originate in ductal tissue and approximately 10\% are lobular. Histological subtypes of breast cancer are summarised in Table 2.

#### Table 2: Breast cancer subtypes

<table>
<thead>
<tr>
<th>Type of Malignancy</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In Situ Disease</strong></td>
<td></td>
</tr>
<tr>
<td>Ductal Carcinoma in Situ (DCIS)</td>
<td></td>
</tr>
<tr>
<td>Lobular Carcinoma in Situ (LCIS)</td>
<td></td>
</tr>
<tr>
<td><strong>Invasive Cancer</strong></td>
<td>&gt;80%</td>
</tr>
<tr>
<td>Invasive Ductal Carcinoma (IDC)</td>
<td>10%</td>
</tr>
<tr>
<td>Invasive Lobular Carcinoma (ILC)</td>
<td>1-4%</td>
</tr>
<tr>
<td>Inflammatory breast cancer</td>
<td>1-4%</td>
</tr>
<tr>
<td>Medullary Breast Cancers</td>
<td>2%</td>
</tr>
<tr>
<td>Mucinous Breast Cancer</td>
<td>1%</td>
</tr>
<tr>
<td>Tubular Breast Cancer</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Angiosarcoma</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Basal type</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Phylloides</td>
<td>&lt;1%</td>
</tr>
<tr>
<td>Papillary</td>
<td>&lt;1%</td>
</tr>
</tbody>
</table>

Management of breast cancer is based around surgical excision +/- (neo) adjuvant chemotherapy, radiotherapy or endocrine therapy depending on tumour characteristics and lymphatic involvement. Breast cancer is a highly heterogeneous disease with disparate patient outcomes dependent on
multiple clinical factors. Traditional predictors of breast cancer survival include tumour stage, grade, age (5y survival is lower for younger women due to a higher incidence of more aggressive tumours), hormone receptor status, oncogene expression and cell type. In 2009, nodal status, tumour size, histological grade and type, peritumoural vascular invasion, HER2 and hormone receptor status were categorized in a consensus statement as the most useful markers in clinical patient management\(^8^4\). The average 15y survival for women with a diagnosis of Grade I cancer is 93.5% but is significantly decreased at 65.3% for those with grade III disease. Breast cancer is also staged using the TNM system (see Appendix Table 25). In addition to this, tumours are graded G1-G3 based on their level of histological differentiation and, like colorectal cancer, TNM classifications are combined to produce a tumour stage (see Appendix Table 26) which reflects prognosis and influences the clinical management strategy.

### 1.7.1 Hormone Receptors

Hormone receptors (oestrogen and progesterone) are proteins expressed on breast cancer cells, which stimulate proliferation, increased cell division, DNA replication and therefore tumour growth in response to hormone stimulation. Approximately 80% of breast cancers are oestrogen-receptor (ER) positive, 65% are ER and PR receptor positive, 13% are ER positive and PR negative, 2% of breast cancers are ER negative and PR positive and 25% of all breast cancers are considered to be hormone receptor negative\(^8^5\). Oestrogen receptors are required for oestrogen stimulated growth and proliferation and receptor positivity indicates that the disease is likely to respond to adjuvant endocrine therapies, which either reduce circulating levels of oestrogen, such as such as luteinising hormone releasing hormone agonists and aromatase inhibitors, or competitively bind and antagonise oestrogen receptors, such as tamoxifen\(^8^6\). There have been some difficulties surrounding definitions of oestrogen receptor status in breast cancer including defining the cut off point at which a tumour is considered to be oestrogen receptor negative and the patient will therefore see no benefit from such adjuvant therapy, and identifying a reliable, quantitative assay with adequate concordance between laboratories. Measurement of ER status was originally by biochemical dextrose coated charcoal (DCC)
techniques, which contributed much of the clinical data used to correlate ER expression with response to tamoxifen and clinical outcome. Monoclonal antibodies to ER were then developed in the 1980s, which allowed less laborious and more precise clinical testing, however, cut off values were based upon those that correlated with DCC. Immunohistochemistry (IHC) methods were developed in the 1990s and remain the most common type of assay used today, using cut off values that pertain to previous clinical data based upon ER receptor status results as obtained by DCC. Current guidelines suggest an IHC cutoff of 1% to define a tumor as ER positive. Few truly ER positive tumours have IHC staining under 10%, however initial ER testing concordance between laboratories has been shown to be as low as 85%, due to pre- analytical and analytical reproducibility and standardisation issues. False negative results can still occur using an IHC assay and can lead to emission of endocrine therapy that could potentially improve prognosis. Setting the cut off at this very low level reduces the risk of false negatives albeit with the resulting risk of an increase in false positive results. Hormone therapy is most commonly used after surgery and any necessary chemotherapy, to reduce the risk of cancer recurrence but can be used pre-operatively to downstage large or locally advanced cancers.

Human Epidermal Growth Factor Receptor-2 (HER2) is a protein normally expressed on the surface of breast epithelial cells. The HER2 protein is a type I transmembrane growth factor receptor. In response to extracellular signals, ligand binding to the extracellular domain activates intracellular signaling pathways, via dimerization and transphosphorylation of the intracellular domain, to activate downstream second messenger pathways and other transmembrane signaling pathways with subsequent diverse intracellular physiological effects. HER2 protein overexpression is present in approximately 25% of breast cancers and is associated with an aggressive tumour phenotype and poor prognosis, however these tumours do respond to the addition of Trastuzumab, HER2 receptor antagonist, into the treatment regime which leads to an increase in disease free and overall survival in this patient cohort. Breast cancers can have up to 100 fold increase in HER2 protein expression leading to the expression of up to 2 million receptors at the tumour cell surface. Currently, immunohistochemistry is used to determine levels of HER2 protein within clinical laboratories, with the addition of fluorescence in-situ hybridisation when results are equivocal. Immunohistochemistry and
corresponding laboratory results are shown in table 3.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Expression of HER2 protein</th>
<th>Laboratory result</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–1+</td>
<td>Normal amount of HER2</td>
<td>HER2 negative</td>
</tr>
<tr>
<td>2+</td>
<td>Moderate amount of HER2</td>
<td>Equivocal</td>
</tr>
<tr>
<td>3+</td>
<td>Higher than normal level of HER2.</td>
<td>HER2 positive.</td>
</tr>
</tbody>
</table>

1.7.2 Surgical Management

Surgery plays a significant role in the management of all but the most advanced cases of breast cancer. Surgical approaches must take into account both management of the primary breast lesion and potential metastases to the axilla. Surgical options for tumour excision range from wide local excision, where only the tumour and a small volume rim of surrounding tissue are excised, quadrantectomy where the affected quadrant is removed, mastectomy and skin-sparing mastectomy with or without immediate or delayed reconstruction. Choice of surgical approach is based upon multiple clinical and oncological factors, potential cosmetic outcome and patient preference. In the UK, the majority of patients undergoing surgery for breast cancer have a breast conserving procedure, where only the tumour and a small proportion of surrounding normal breast are removed\(^9^4\). For the majority of women this is more cosmetically acceptable and, when combined with postoperative radiotherapy, produces similar survival outcomes to mastectomy\(^9^4\). Success of breast conserving surgery relies upon adequate resection of all tumour tissue, however, 20% of women who had primary breast conserving surgery in the UK between 2005 and 2008 required re-operation within three months due to positive resection margins, and one in seven required further surgery after that. There are currently limited methods to determine which patients will be at high risk for positive resection margins, although reoperation is more common among women with a carcinoma in situ component recorded at the time of primary surgery than in those without. It is also more common among younger women\(^9^5\). Current strategies for the intra-operative detection of positive resection margins include intra-operative frozen section, touch imprint cytology and specimen radiography.
1.7.3 Adjuvant Therapy

ER/PR and HER 2 receptor expression are strongly predictive biomarkers of response to directed, targeted therapies. Results of these tests alone are key determinants of which adjuvant treatment the patient receives. Few oestrogen receptor positive and/or lymph node negative patients receive chemotherapy, whereas ER negative and LN positive patients do and HER2 negative patients do not receive trastuzumab. Although a proportion of breast cancers will benefit from systemic chemotherapy, endocrine and HER-2 based therapies, traditional pathological prognostic categories are not currently sufficient to fully identify which patients will see the greatest benefit from these treatments. Despite the development of successful adjuvant therapies for each patient cohort, a significant proportion of patients with breast cancer still ultimately relapse with metastatic disease, whilst some will undergo intensive chemotherapy which may not actually be required, either because they have a low risk of relapse or because they exhibit inherent tumour biology that is less sensitive to chemotherapy and more responsive to early endocrine therapy\textsuperscript{96}. Recently improved understanding of the biology of breast cancer has led to the identification of novel molecular targets such as the phosphoinositide 3 kinase (PI3K)/Akt/mTOR pathway\textsuperscript{97}, (associated with endocrine resistance) and poly(ADP-ribose) polymerase (a possible molecular target in triple negative breast cancer) that are amenable to therapeutic intervention. Recent genomic work has suggested that many more molecular subgroups of breast cancer may exist than are currently defined, which may be amenable to more personalized management strategies\textsuperscript{98}. There is a need for alternative methods of risk stratification of breast cancer patients within the clinical environment, and significant progress has been made in the development of tissue based tests which aim to predict risk of recurrence in patients with early stage cancer, and in some cases DCIS. These include gene expression profiling assays such as Oncotype Dx, MammaPrint, PAM50 and expanded immunohistochemistry tests such as Immunohistochemistry (ICH)\textsuperscript{4}, a score based on ER, PR, HER2 and Ki67 immunochemistry and Mammostrat based on SLC7A5, HTF9C, P53, NDRG1, and CEACAM5. The resulting recurrence scores aid clinical decision making around the benefit of adjuvant chemotherapy or radiotherapy in those patients in whom the risk vs. benefit of adjuvant therapy is not well defined. Oncotype Dx is the only one of these tests presently recommended by NICE.
for use in clinical practice as a tool to discriminate those with ER positive, lymph node negative, HER2 negative early breast cancer who will benefit from adjuvant chemotherapy. Technologies to provide the surgeon with robust, objective real time data on the biochemical or molecular status of tissues during clinical and surgical management of breast cancer could improve the efficacy of management. Improved understanding of the molecular and biological basis for the disease has already lead to an increased interest in the tailoring of medical and oncological therapies to conform to the individual needs of each patient, and the availability of individualised, molecular tissue information could further contribute to provision of personalized management of breast cancer.

In summary, precision surgical technologies such as the iKnife offer significant potential to improve the quality of the surgical management of both early and late disease in both breast and colorectal cancer. The literature reveals a great heterogeneity in the current management of both conditions and an unmet need for improved tumour phenotyping and better patient stratification if patient outcomes from cancer treatment in general are to be improved.

1.8 Lipid Metabolism in Cancer

The most well known metabolic change in cancer cells is the Warburg effect\textsuperscript{99}. In contrast to normal differentiated cells, tumour cells rely heavily on aerobic glycolysis for the generation of the energy required for cellular processes, rather than the more efficient mitochondrial oxidative phosphorylation. Rapidly growing tumour cells may have glycolytic rates significantly higher than their tissues of origin, even in the presence of normal oxygen levels, requiring a high rate of glucose uptake to meet energy requirements and support tumour progression. It has been more recently proposed that the switch to aerobic glycolysis serves as a mechanism to provide cancer cells with not only energy but also the building blocks for macromolecule synthesis, such as carbohydrates, proteins, and lipids\textsuperscript{100}.

In addition to the Warburg effect, alteration in lipid metabolism is increasingly being recognised as an important distinctive feature of cancer cells\textsuperscript{101}. Cells acquire fatty acids to meet their metabolic demand from two major sources; exogenously derived (dietary) fatty acids and \textit{de novo} endogenous synthesis.
Most adult normal cells preferentially utilise exogenous fatty acids, with the exception of the lactating breast and cycling endometrium where biosynthesis of endogenous de novo fatty acids is catalysed by the enzyme fatty acid synthase (FASN). Other human tissues acquire the majority of required fatty acids from the circulation, and accordingly de-novo lipogenesis and expression of lipogenic enzymes, such as FASN, are low\(^\text{102}\). In contrast, cancer cells preferentially endogenously synthesize fatty acids, even in an abundance of extracellular fatty acids\(^\text{103}\), in fact, de novo synthesis may account for more than 93% of triacylglycerol fatty acids in tumour cells\(^\text{104}\). The relationship between this neoplastic lipogenesis and the tumour associated “glycotic switch” remains poorly understood\(^\text{104}\). The lipogenic phenotype almost certainly conveys a growth advantage, with dysregulated lipid metabolism and associated signaling pathways altered to meet the abnormal demand for cancer cell proliferation and survival. Continuous de novo lipogenesis provides cancer cells with membrane building blocks, signaling lipid molecules and energy to support rapid cell proliferation, however the implications of increased fatty acid synthesis for tumour biology and cancer cell proliferation remain largely unknown. Increased fatty acid synthesis has been assumed to lead to the upregulation of phospholipid synthesis that meets the increased need for membrane production in highly proliferative cancer cells. These phospholipids facilitate signal transduction, intracellular trafficking and polarization as well as providing essential structural lipids\(^\text{104,105}\). It is proposed that, in epithelial cancer cells, FASN is primarily involved in production of phospholipids, which are incorporated into detergent resistant membrane microdomains or rafts\(^\text{105}\). This is because the main effect of FASN modulation seen in cancer cells was a reduction in the synthesis of major phosphatidyl phospholipids, primarily phosphatidylcholine (PC) followed by phosphatidylethanolamine (PE), phosphatidylserine (PS) and phosphatidylinositol (PI)\(^\text{105,106}\). In addition, choline kinase, the first enzyme in the Kennedy pathway responsible for the de novo synthesis of phosphatidylcholine, which in addition to phosphatidylethanolamine is the most abundant phospholipid in the cell membrane, also appears to be markedly overexpressed in many cancers\(^\text{107,108}\).

Human cells are unable to synthesize polyunsaturated and highly-unsaturated fatty acids from the fully saturated precursors produced by FASN, and as such the newly synthesized phospholipids are mainly saturated and mono-unsaturated (short chain fatty acids (SCFAs) are converted into monounsaturated
fatty acids by the rate limiting action of the enzyme SCD1 and the resulting products converted by a series of enzymes to generated monounsaturated fatty acids\textsuperscript{109}). Diet derived lipids do contain polyunsaturated fatty acids and \textit{de novo} lipogenesis in cancer cells may not only lead to quantitative changes in membrane lipids required for cell growth but also compositional change\textsuperscript{105}. The alteration of lipid metabolism in cancer functions downstream from oncogenic signaling. The expression and upregulation of enzymes involved in fatty acid synthesis, start at an early stage and, in addition to increased levels of saturated phosphatidylcholine, have been associated with aggressive tumour type, risk of recurrence and poor prognosis in breast cancer\textsuperscript{110,111} and reduced disease free survival in colorectal cancer\textsuperscript{112}. As such FASN is a potential target both for prognostic stratification and therapeutic intervention and alterations in structural membrane phospholipids may be an important feature of cancer cell metabolism. FASN is highly expressed in both invasive and \textit{in situ} breast cancer\textsuperscript{103,113} and level of expression also confers prognostic information. In breast cancer, high levels of endogenous FASN activity have been associated with poor prognosis and higher risk of recurrence\textsuperscript{105}. In addition to this, HER2 overexpression in breast cancer, confers a lipogenic phenotype in addition to a more aggressive one\textsuperscript{110}. The proposed oncogenic mechanism is that HER2 overexpression increases the translation of FASN, which in turn increases HER2 signaling, resulting in uncontrolled cell growth\textsuperscript{114}. FASN phosphorylation by HER2 has therefore become a potential therapeutic target in HER2 positive breast cancer\textsuperscript{114}. FASN also leads to inhibition of the intrinsic pathway of cellular apoptosis\textsuperscript{115} and inhibition of fatty acid synthesis has been shown to induce programmed cell death in human breast cancer cells\textsuperscript{116}.

The proposed downstream effects of FASN on membrane phospholipids in breast cancer cells have been extensively investigated by mass spectrometry and NMR based lipidomic analyses. A consistent finding has been significant differences in levels of choline phospholipid metabolites in malignant breast tissue compared to normal epithelial cells, particularly elevation of phosphocholine and total choline-containing metabolites (The sum of choline (Cho), phosphocholine (PCho) and glycerophosphocholine (GPC))\textsuperscript{117,118}. In addition to higher levels of choline containing metabolites, there
appears to be a GPC to PCho switch in malignant cells, i.e. PCho as opposed to GPC becomes the major choline metabolite leading to a significantly higher PC/GPC ratio in breast cancer cells as compared to normal tissue\textsuperscript{119}. Phosphocholine together with other phospholipids such as phosphatidylethanolamine and neutral lipids, forms the characteristic lipid bilayer structure of cellular membranes\textsuperscript{120}. Responsible underlying molecular pathways are reported to be a combination of high rates of choline phosphorylation due to overexpression of choline kinase and increased breakdown of membrane phosphatidylcholine secondary to increased phospholipase C and D activity in cancer cells\textsuperscript{107}. These pathways have been validated in multiple studies\textsuperscript{121,108,122}. NMR analysis of intact tissue samples or tissue sample extracts has been the most valuable in providing insight into the choline metabolite changes associated with malignant transformation in breast tissue. Metabolite changes, choline metabolites highlighted, in breast cancer tissue and normal healthy tissue as measured by NMR analysis of intact breast tissue (MAS NMR) and tissue extract (\textsuperscript{1}H NMR) are summarized in Table 4. The increased PCho level in malignant cells was initially attributed to higher rates of cell proliferation. However, non malignant breast and prostate cells induced to proliferate as rapidly as cancer cells still exhibited lower levels of PCho and total choline than cancer cells\textsuperscript{119} and therefore malignant transformation, rather than simply cell proliferation, may be contributing to the abnormal phospholipid metabolism seen in cancer cells\textsuperscript{123}.

Table 4: MAS NMR/\textsuperscript{1}H NMR studies in breast cancer tissue/extract demonstrating metabolite alterations in cancer tissue compared to healthy tissue and measurable metabolite changes associated with individual oncological prognostic features.

<table>
<thead>
<tr>
<th>Study/Modality</th>
<th>Normal Breast Tissue</th>
<th>Breast Cancer Tissue</th>
<th>Clinical Features</th>
</tr>
</thead>
</table>
| Cheng et al 1998\textsuperscript{124} MAS NMR | ↓Phosphocholine, ↑Choline | ↑Phosphocholine | Higher grade
↑PCho: Ch ratio
↑lactate:choline ratio |
| Beckonert et al 2002\textsuperscript{125} Tissue extract \textsuperscript{1}HNMR | ↑Myoinositol | ↑Phosphocholine (higher in higher grade)
↓Glycerophosphocholine
↑Taurine
↓Glucose | Higher grade
↑Phosphocholine,
↑phosphoethanolamine,
↑Alanine, ↑Taurine, ↑fatty acids
↓Inositol, ↓glucose, ↑UDP-hexose. |
| Sitter et al 2006\textsuperscript{126} MAS NMR | ↑Glycerophosphocholine Ratio of PC:Cho high | ↑Phosphocholine | Tumour >2cm
↑Choline
↑Glycine |
<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Methodology</th>
<th>Findings</th>
</tr>
</thead>
</table>
| Bathen et al 2007            | 2007 | MAS NMR     | Lymph node involvement 
- ↑glycine 
- ↑taurine |
| Sitter et al 2010            | 2010 | MAS NMR     | Poor prognosis 
- ↑Glycine 
- Altered ratio of Taurine:glycine 
- GPC:glycine 
- Total choline: glycine |
| Li et al 2011                | 2011 | MAS NMR     | Lymph node involvement 
- ↑Phosphocholine 
- ↑ Taurine |
| Cao et al 2011               | 2011 | MAS NMR     | Survivors 
- ↑tCho and ↓lactate pretreatment compared to non survivors pretreatment. 
- Survivors ↓GPC, ↓glycine 
- ↓lactate post treatment compared to non survivors post treatment 
- ↑taurine post treatment best predictor of clinical response. |
| Cao et al 2012               | 2012 | MAS NMR     | Non survivors 
- ↓GPC and taurine post NAC compared to pre NAC. 
- ↑ lactate and glycine 
- Survivors ↓PCh and choline containing compounds, ↓ glycine and ↑ glucose after NAC compared to pre. |
| Giskeodegard et al 2012      | 2012 | MAS NMR     | Poor prognosis 
- ↑phosphocholine 
- ↑lactate, ↑glycine, 
- ↓glucose, ↓creatine, ↓taurine |
| Bathen et al 2013            | 2013 | MAS NMR     | Triple negative 
- ↑ choline, ↑ GPC , ↓ creatine. ↓ glutamine, ↑ glutamate 
- ER negative 
- ↑glycine, ↑choline, ↑lactate |
| Cao et al 2014               | 2014 | MAS NMR     | |
Overexpression of FASN is also a common feature of colorectal cancer\textsuperscript{134} and high levels have been associated with poor prognostic indicators\textsuperscript{135} and tumour location within the left side of the colon. Similarly, altered choline metabolism has become a metabolic hallmark of colorectal cancer oncogenesis and tumour progression\textsuperscript{136} secondary to changes in membrane phospholipid metabolism and overexpression of choline kinase. Metabolite changes, measured by MAS NMR of intact tissues and NMR analysis of tissue extracts, secondary to malignant change are summarised in table 5.

<table>
<thead>
<tr>
<th>Study</th>
<th>Benign Colorectal Tissue</th>
<th>Malignant Colorectal Tissue</th>
<th>Clinical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merchant et al 1991</td>
<td>31P MRS</td>
<td>↑PCho,</td>
<td>↑lysophosphatidylcholine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↓phosphatidylethanolamine,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>↓sphingomyelin</td>
<td></td>
</tr>
<tr>
<td>Moreno et al 1993</td>
<td>1H NMR</td>
<td>↑taurine,</td>
<td>↓polyethyleneglycol (reduced absorption)</td>
</tr>
<tr>
<td>Chan et al 2009</td>
<td>MAS NMR + GCMS</td>
<td>↑Lipids,</td>
<td>↑choline-containing compounds (ChoCC), ↑taurine,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑polyethylene glycol,</td>
<td>↑scylo-inositol, ↑glycine,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑glucose,</td>
<td>↑phosphoethanolamine (PE),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑fumarate,</td>
<td>↑lactate and ↑phosphocholine (PC),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑malate,</td>
<td>↑2-Hydroxy-3-methylvalerate,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑mannose,</td>
<td>↑l-1-Proline,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑galactose,</td>
<td>↑L-phenylalanine,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑1-hexadecanol and</td>
<td>↑fatty acids (Palmitic acid, Marganic acid, Oleic acid, Stearic acid, 11,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑arachidonic acid</td>
<td>14-eicosadienoic acid, 11-eicosadienoic acid, ) ↑ uridine,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑1-O-Heptadecylycerol,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑Propyl octadecanoate,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑Cholesterol.</td>
</tr>
<tr>
<td>Righi et al 2009</td>
<td></td>
<td>↑Choline containing</td>
<td>↑Taurine,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>compounds, ↑Creatine,</td>
<td>↑Acetate,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>↑scylo-inositol,</td>
<td>↑Lactate,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>↑lipids</td>
</tr>
<tr>
<td>MAS NMR</td>
<td>↑Glutamate, ↑myo-inositol, and ↑Glucose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang et al 2013\textsuperscript{140}</td>
<td>↑lactate, ↑threonine, ↑acetate, ↑glutathione, ↑uracil, ↑succinate, ↑serine, ↑formate, ↑lysine and ↑tyrosine, ↓myo-inositol, ↓taurine, ↓phosphocreatine, ↓creatine, ↓betaine and ↓dimethylglycine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>¹H NMR Rectal cancer only</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jiménez et al 2013\textsuperscript{141}</td>
<td>↓lips (triglycerides), ↑isoglutamine, ↑tyrosine, ↑Cho, ↑PCho, ↑taurine, ↑lactate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAS NMR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mirnezami et al 2014\textsuperscript{142}</td>
<td>↑lactate, ↑taurine, ↑isoglutamine, ↓lips/triglycerides.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon cancer</td>
<td>↑taurine associated with increasing T stage, ↑valine, ↑alanine, ↑phenylalanine and ↑tyrosine in T4 tumours.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lymphnode involvement</td>
<td>↑creatine, ↑scyllo-inositol</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Imaging mass spectrometry studies have demonstrated histological specificity of certain lipids\textsuperscript{143} and identification of individual tissue types based on the presence of these lipid markers\textsuperscript{144,145,146}. Recently, MALDI MSI has been applied to investigate the effects of de novo lipogenesis on 6 common cancer types including breast and colorectal cancer. Findings suggest significantly increased levels of cancer related monounsaturated fatty acids and monounsaturated phosphatidylcholines relative to polyunsaturated fatty acids and polyunsaturated phosphatidylcholines in the cancer microenvironment compared to adjacent normal tissue associated with de novo lipogenesis, with increased levels of FASN, CKα and SCD1, further consolidating evidence of altered lipid and choline metabolism in the oncological process\textsuperscript{109}.

### 1.9 Metabonomics Studies in Colorectal Cancer

There have been recent attempts to develop novel omics based approaches for use in the risk stratification of colorectal cancer. The Oncotype Dx assay, similar to that used in breast cancer, was developed from mRNA expression profiles of 18 target genes (7 for relapse free survival in colon cancer,
6 to predict response to 5-fluorouracil/leucovorin chemotherapy and 5 reference genes) using RT PCR of resected tissue samples, to create a score to predict CRC recurrence in stage II cancers managed surgically, for whom the indication for neoadjuvant therapy is unclear. The 18 gene panel was found to be a valid predictor of disease free survival but not of treatment response, and therefore only the genes for disease free survival were included in the marketed test\textsuperscript{147}. Individuals are assigned a recurrence score (low, intermediate or high) which is associated with recurrence risk, shorter disease free survival and overall survival but does not predict for benefit from 5-FU chemotherapy in these patients. Mass Spectrometry based proteomics studies have identified potential serum biomarkers for diagnosis of colorectal cancer\textsuperscript{80}, candidate serum biomarkers with detectable alteration during chemotherapy in accordance with patient response to treatment\textsuperscript{148}, and potential serum biomarkers of lymphnode metastasis\textsuperscript{149}. Predose serum $^1$H NMR metabolic profiles have been shows to predict a higher risk of developing side effects secondary to toxicity in patients with locally advanced or metastatic colorectal adenocarcinoma being treated with the oral chemotherapeutic agent capecitabine, a pro-drug of 5-fluorouracil\textsuperscript{150}. Metabonomics studies of colorectal cancer tissue have also proved promising in the identification of colorectal malignancy, prediction of tumour characteristics and have provided insight into the biochemistry of colorectal carcinogenesis. For example, the detection of a tumour associated field change or microenvironment by MALDI analysis\textsuperscript{151} has implications for both furthering understanding of tumour biology and oncogenesis and also in obtaining tumour free margins in cancer resection surgery. MAS NMR has been shown to differentiate between colorectal adenocarcinoma and normal adjacent mucosa with an AUC of 0.98, and the authors have described some key metabolic differences between the 2 tissue types\textsuperscript{138}. Differences in the metabolic profiles of colonic and rectal adenocarcinoma and, crucially, between T1/2, T3 and T4 tumours have been demonstrated and metabolic differences contributing to the differentiation have been described\textsuperscript{138}. DESI-MSI has been shown to differentiate between normal colonic mucosa, smooth muscle tissue and adenocarcinoma with a classification accuracy of 95%. KRAS mutation status could also be correctly predicted 86.7% of the time\textsuperscript{142}, suggesting these methods may present faster, cheaper alternatives to expensive genetic analysis, as pre-screening methods. The use of these approaches to improve stratification of disease
and prediction of response to therapy, particularly in the pre-operative stage, could impact clinical decision making in the management of colorectal cancer.

1.10 Metabonomics Studies in Breast Cancer

The ability of metabonomic strategies to track downstream cellular changes facilitates deeper understanding of the biochemical processes underlying breast cancer aetiopathology and response to therapy. Over the last 2 decades, such strategies have been employed in breast cancer research in the development of molecular based tissue diagnostics\textsuperscript{152,92,153} biofluid biomarker discovery\textsuperscript{154,155} and preliminary prediction of metastasis, recurrence\textsuperscript{156}, prognosis and outcome\textsuperscript{157}. More than 30 metabolites in breast cancer tissue have been identified and assigned using MAS NMR of intact breast tissue and extract from breast cancer tissue\textsuperscript{158}. Levels of metabolites including choline and its metabolites, taurine, glycine, lactate, myo-inositol and UDP hexose, as well as relative intensities of glycerophosphocholine (GPC), phosphocholine (PCho) and choline have been shown to be relevant in the differentiation of malignant and non malignant breast tissue\textsuperscript{124}. Similar shifts in metabolite profiles (relative intensities of taurine, GPC, lactate, glycine, scyllo-inositol and PCho) have been demonstrated in the differentiation of node positive disease\textsuperscript{127} and increased levels of lactate and glycine have been shown to be associated with a reduced 5 year survival in oestrogen receptor positive cancers\textsuperscript{130}. Using DESI MSI, metabolic profiles of breast cancer tissue were sufficient for the accurate differentiation of grade 2 and grade 3, and ER positive and negative cancers (AUC 0.97 and 0.96 respectively)\textsuperscript{159} and a further high-performance-liquid-chromatography–electrospray-MS study of breast cancer tissue extracts demonstrated that increasing tumour grade and negative HR status are both individually correlated with an increase in concentration of the same lipids\textsuperscript{110}.

1.11 Intra-Operative Tissue Diagnostics

In situations where surgical margins may be compromised, the most commonly used method for tissue identification is the intra-operative frozen section; the current gold standard for intra-operative
determination of resection margins, lymph node status and tissue identification. The frozen section process involves sending the resected tissue to the pathology lab where it is frozen in a cryostat, cut with a microtome, stained and examined by a pathologist. The result is then relayed back to the surgical team and the procedure continues accordingly.

This process, however, has several limitations:

1. It is slow: It takes approximately 30 minutes for the tissue to be sent to the pathology lab, examined and the results to be relayed to the operating theatre team.

2. It places a burden on histopathology services, as a trained pathologist must be on site to examine the prepared tissue section.

3. The process can be technically demanding (both preparation and interpretation) especially under pressure of time.

4. Quality of the slides produced by frozen section is morphologically inferior to standard formalin fixed, wax embedded tissue processing and therefore fixed tissue processing is preferred for more accurate diagnosis.

5. The number of points that can reasonably be sampled during a single surgical resection is limited and so frozen section must be targeted to particular areas of concern which can be difficult to identify in, for example, a wide local excision for breast cancer.

6. The process is expensive with an average cost of $300.

With recent advances in spectroscopic techniques, a multitude of spectroscopic methods for the identification of individual tissue morphologies have been developed, aimed at providing the clinician with histology level, tissue specific information, quickly and accurately within the operating theatre environment.

The ideal modality for intra-operative tissue diagnostics would possess the following characteristics:

- Rapid
- High throughput
- Objective
- More accurate than currently available methods (frozen section pathology or touch imprint cytology)
- Non destructive
- Require no tissue pre-processing
- Cost effective
- Safe e.g. require no exposure to ionising radiation

The majority of significant advances in technologies for intra-operative margin detection have been made in the field of breast cancer, particularly since the introduction of breast conserving surgery as the preferred surgical approach for the majority of patients. These include techniques for examining the excised specimen or the surgical cavity with the aim of reducing positive margins and therefore need for re-excision. Traditional pathological methods, such as frozen section and imprint cytology perform well in terms of sensitivity and specificity but significantly increase operating time for limited sampling points\textsuperscript{162}. More novel methods such as digital specimen mammography, radiofrequency spectroscopy, and intra-operative ultrasound were found to be significantly more time efficient in a recent systematic review, but suffered from poor accuracy or had an incomplete evidence base\textsuperscript{162}. A summary of technologies, both in use and in development, for intraoperative margin assessment in breast cancer is summarized in table 6. The MarginProbe, a new radiofrequency spectroscopy probe performs well in terms of concordance with formal histopathology\textsuperscript{162}. Spectroscopic techniques are attractive because they provide reproducible, quantitative biochemical information about the morphological nature of tissues non-destructively, non-invasively and without the use of ionising radiation. The MarginProbe is a commercially available example of radiofrequency spectroscopy designed for intra-operative use. The device was initially developed for use in breast conserving cancer resections and has been further developed for potential use in prostate cancer and other solid organ malignancies. The device measures the differing electromagnetic signature reflected by healthy and cancerous tissue, a result of the inherent structural differences between normal and malignant cells. The system is comprised of a single-use probe and a portable console and the technique is rapid (<3s) and non destructive. Approved by the FDA in January 2013, the device is currently commercially
available in Europe and Israel. When used as adjunctive tool to standard surgical care in 22 patients with impalpable ductal carcinoma in situ undergoing breast conserving surgery, use of the probe reduced re-excision rates from 38.8% to 18%\textsuperscript{163}. A prospective randomised controlled trial with 300 recruited patients demonstrated an increase in the intra-operative identification and re-excision of all positive margins in the device group (60% vs. 41%). Repeat lumpectomy rate was also reduced in the device group compared to control group (5.6% vs. 12.7%)\textsuperscript{164}. A number of alternative spectroscopic techniques are in development including optical spectroscopy, Fourier transform infrared spectroscopy and fluorescence spectroscopy. Optical spectroscopy, including diffuse reflectance spectroscopy, utilises the interaction of light with tissue at the cellular and subcellular level to provide a rapid, in vivo diagnosis based upon differing absorption, reflectance and scattering spectra produced by differences in cell morphology associated with malignant transformation\textsuperscript{165}. A single optical probe has been developed for use in skin lesions, which has potential as a tool for identification of malignant and benign tissue in vivo\textsuperscript{166}. A similar single-fiber optical probe system has been described for the intraoperative detection of low and high-grade brain tumors and identification of positive tumour margins\textsuperscript{166}. Diffuse Reflectance spectroscopy has been applied to the assessment of surgical margins during breast conserving surgery using β-carotene/scattering ratios. Positive margins were identified with a sensitivity and specificity of 79% and 67% respectively\textsuperscript{167}. Fluorescence spectroscopy uses either native fluorescence from endogenous fluorophores, such as flavins, elastin and collagen, or emission from exogenously administered fluorescent substances for in-vivo quantitative tissue diagnostics. Advances in light delivery and collection systems (particularly fiberoptic systems) have facilitated the growth of fluorescence-based techniques as potentially translatable tools for the clinical environment. A near-infrared capable laparoscopic tool has been optimally adapted to the fluorescence characteristics of indocyanine green (ICG), and neuro-endoscopes to allow endoscopically assisted 5-ALA-fluorescence guided resections to be performed during malignant brain tumour surgery.
Table 6: Current techniques for use in the intra-operative assessment of breast cancer resection margins.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Study</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>Increased operating time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intraoperative DSM</td>
<td>Kaufman et al 2007</td>
<td>36%</td>
<td>71%</td>
<td>Not reported</td>
<td>-19 mins</td>
</tr>
<tr>
<td></td>
<td>Kim et al 2016</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>-2 min</td>
</tr>
<tr>
<td>Intraoperative ultrasound</td>
<td>Moore et al 2001</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>-15 mins</td>
</tr>
<tr>
<td></td>
<td>Rahusen et al 2002</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>-1 min</td>
</tr>
<tr>
<td></td>
<td>James et al 2009</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Olsha et al 2011</td>
<td>100%</td>
<td>74%</td>
<td>99.6%</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Doyle et al 2011</td>
<td>25%</td>
<td>95%</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Ramos et al 2013</td>
<td>80%</td>
<td>86.6%</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>RF spectroscopy probe</td>
<td>Karne et al 2007</td>
<td>71%</td>
<td>68%</td>
<td>Not reported</td>
<td>+7.37 mins</td>
</tr>
<tr>
<td></td>
<td>Allweiss et al 2008</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Thill et al 2011</td>
<td>Not reported</td>
<td>Not reported</td>
<td>73%</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Rivera et al 2012</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>Intraoperative frozen section</td>
<td>Olson et al 2007</td>
<td>73%</td>
<td>99%</td>
<td>98%</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Weber et al 2008</td>
<td>80%</td>
<td>87.5%</td>
<td>83.8%</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Dener et al 2009</td>
<td>Not reported</td>
<td>Not reported</td>
<td>94%</td>
<td>+25 mins</td>
</tr>
<tr>
<td>Procedure</td>
<td>Study</td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Min. Time</td>
<td>Notes</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------------------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>Touch imprint cytology</td>
<td>Creager et al 2002&lt;sup&gt;183&lt;/sup&gt;</td>
<td>80%</td>
<td>85%</td>
<td>+20 mins</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Martin et al 2011&lt;sup&gt;184&lt;/sup&gt;</td>
<td>70%</td>
<td>97.1%</td>
<td>Not reported</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sumiyoshi et al 2011</td>
<td>85%</td>
<td>100%</td>
<td>+25 mins</td>
<td></td>
</tr>
<tr>
<td>Optical Coherence Tomography</td>
<td>Nguyen et al 2016&lt;sup&gt;185&lt;/sup&gt;</td>
<td>100%</td>
<td>82%</td>
<td>90%</td>
<td>Not reported</td>
</tr>
<tr>
<td>Gamma camera</td>
<td>Duarte et al 2007&lt;sup&gt;186&lt;/sup&gt;</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Paredes et al 2008&lt;sup&gt;187&lt;/sup&gt;</td>
<td>Not reported</td>
<td>Not reported</td>
<td>60%</td>
<td>+5 mins</td>
</tr>
<tr>
<td>Gross tissue inspection/specimen radiography</td>
<td>Chagpar et al 2003&lt;sup&gt;188&lt;/sup&gt;</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Fleming et al 2004&lt;sup&gt;189&lt;/sup&gt;</td>
<td>73%</td>
<td>88%</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td></td>
<td>Cabioglu et al 2007&lt;sup&gt;190&lt;/sup&gt;</td>
<td>92%</td>
<td>78%</td>
<td>87%</td>
<td>Not reported</td>
</tr>
<tr>
<td>2 view specimen mammography</td>
<td>McCormick et al 2004&lt;sup&gt;191&lt;/sup&gt;</td>
<td>55%</td>
<td>88%</td>
<td>Not reported</td>
<td>+15 mins</td>
</tr>
</tbody>
</table>

In addition to tools for intra-operative tissue identification, there has been significant recent investment in developing image-based techniques for endoscopic diagnostic use for the purposes of identifying dysplastic lesions *in vivo*, including Confocal Laser Endomicroscopy (CLE), Magnified chromoendoscopy and Narrow Band Imaging (NBI). Confocal Laser Endomicroscopy is used to obtain very high-resolution images of the mucosa and can be based on tissue reflectance (no contrast agent required but low resolution) or fluorescence (local or IV contrast agent required). Varying sensitivity and specificity has been reported depending on the size of lesion, experience of operator and quality of image captured. Xie et al reported a sensitivity and specificity of 97.1 % and 100% in the identification...
of adenomatous polyps >10mm in size, but this was reduced to 90.3% and of 95.7%, respectively in polyps under 10mm\textsuperscript{192}. Narrow-band imaging utilises short wavelength, narrow-bandwidth blue light which is issued from the colonoscope, enhancing mucosal detail, particularly vascular structures, allowing assessment of microvascular density via vascular pattern intensity. Neoplastic tissue is characterised by increased angiogenesis, so adenomas appear darker then mucosal tissue when viewed with NBI. A meta-analysis published in 2013\textsuperscript{193} demonstrated an overall sensitivity and specificity of 91% and 82.6% respectively for the identification of neoplastic from non-neoplastic colonic polyps. Probe-based Confocal Laser Endomicroscopy has a reported sensitivity of 86% and specificity of 72% in the detection of dysplastic colonic polyps. However this was based on off-line video footage review and requires administration of IV fluorescein\textsuperscript{194}. Magnified chromoendoscopy which involves the use of 0.5% indigo carmine dye sprayed over the lesion prior to pit pattern and microvessel analysis using magnified colonoscope view, has a reported overall diagnostic accuracy of 95% when performed by a single operator who is expert in magnified colonoscopy.

Despite significant interest and investment in the development of technologies for \textit{in vivo} tissue diagnostics, none of these systems have, as yet, become part of routine clinical practice.
2. Methods Background

2.1 Mass Spectrometry

Mass spectrometry is an analytical technique for the identification of the chemical constituents of a sample by separation of gaseous ions by their mass and charge. The analytical advantages of mass spectrometry over other methods include high sensitivity and speed of analysis. In order for the analysis of gas phase ions to take place, charged particles (ions) must be created from molecules. These ions are then analysed to provide information about the molecular weight and chemical structure of the compound based on their mass to charge \((m/z)\) ratio.

All mass spectrometers therefore consist of three distinct regions.

- Ionisation source
- Mass Analyser
- Detector

**Ionisation Source**

The ionisation source converts analytes of interest into gas phase molecular ions for analysis. Multiple ionisation methods exist including Electron Impact Ionisation (EI), Chemical Ionisation (CI), Electrospray Ionisation (ESI), Fast Atom Bombardment (FAB) and Matrix Assisted Laser Desorption Ionisation (MALDI). Once gas-phase ions have been produced, they can be separated based on their masses, within the mass analyser.

**Mass Analyzer**

The physical property of ions measured by a mass analyser is the mass to charge ratio, rather than the mass alone. Molecular ions and fragment ions are accelerated through the mass spectrometer by manipulation using static or dynamic electric and magnetic fields based on their mass/charge whilst uncharged molecules and fragments are diverted away. There are five principal types of mass analyser in use and these can be divided into two groups: beam analysers and trapping analysers.
Beam analysers transmit ions of different masses successively along a time scale. Ions leave the ion source in a beam and pass through the analyzing field to the detector. They include Time-of-Flight (ToF), Sector (magnetic or electrostatic) and Quadropole instruments. ToF mass analysers separate ions based on their velocity. The ions are formed at the same time in the ion source and then accelerated or pulsed into the ToF drift tube. After acceleration, the ions travel through a fixed distance, typically 0.5–2 metres, before striking the detector. The m/z value allocated to the ion is based on the time it takes to reach the detector. More sophisticated TOF instruments include a reflectron; after travelling through one flight distance, the ions enter an electrostatic mirror that reflects the ions back down a second flight distance to the detector. This increase in distance covered compensates for small differences in the velocities of ions with the same m/z. Quadrupole mass analysers direct ions through four electrodes (rods) that have both AC and DC voltages applied to them. Ions are affected by the electromagnetic field and the alternating voltages affect ion trajectory so that progress down through the quadropole is proportional to their m/z value. Only ions of a certain m/z ratio will reach the detector for a certain range of voltages, the others will have an unstable trajectory and collide with the rods. Using this method a particular range of m/z values can also be selected for analysis, based on voltages applied. Sector mass spectrometers use a static electric or magnetic sector or commonly a combination of the two, as a mass analyser. Ions pass through the magnetic/electric sector in which the magnetic field is applied in a direction perpendicular to the direction of ion motion. This changes the trajectory of the charged particles based on their m/z values.

When trapping analysers are used, the ions are trapped in the analyzing field, after being formed in the analyser itself or being injected from an external ion source, such as electrospray. Trapping analysers include Fourier-transform ion-cyclotron resonance (FT-ICR), Quadropole traps and Orbitraps. The Quadropole ion trap is similar to the Quadropole however, instead of having electric fields in 2 dimensions that allow movement of ions perpendicular to the field, the ion trap has electric fields in 3 dimensions, and so ions become trapped within the field. Ions are then made unstable in a mass selective manner by increasing the voltage applied and are thus detected based on their m/z ratio.
FT-ICR mass analysers determine the m/z ratio of ions based on their fundamental oscillation frequency in a fixed magnetic field. Orbitrap mass analysers trap ions in an orbital motion around a central spindle-like electrode. Harmonic oscillations along the axis of rotation occur at a frequency characteristic of their m/z value.

**Detector**

Once the ions have been filtered by m/z value by the analyser, the detector records the abundance of each m/z value. There are many types of detector, but most work by producing an electronic signal based on the charge or current induced when the surface is when struck by an ion. This is converted into graph which demonstrates the relative current produced by ions of varying m/z ratios.

Mass spectrometry based platforms used in the analysis of biofluids and tissues are summarized in table 7. The most recent developments are mass spectrometry imaging methods for spatially resolved analysis of metabolites in biological tissues. Each ionization method and type of mass analyser possess different features, including the m/z range that can be measured, the mass accuracy, and the achievable resolution. The unit of mass measurement used in mass spectrometry is the unified atomic mass unit (u) or Dalton (Da) which is equal to 1/12th of the mass of an unbound neutral atom of $^{12}$C. The resolution of the instrument is the ability to separate species with almost the same mass and accurately determine the m/z of individual analytes. i.e. when considering ToF MS, it is the power of the instrument to separate ions of similar flight times. The mass accuracy of the instrument refers to the degree of conformity of a measurement to its actual true value and is generally reported as a relative value – parts per million (ppm). Using this measure, the level of variation in Da is dependent on the mass, for example, a mass accuracy of 5 ppm at m/z 300 would allow a variation of +/- 0.0015Da (300 x 5/10^6), but at m/z 3000 would allow a variation of 0.015Da (3000 x5/10^6).
Table 7: Commonly utilised mass spectrometry platforms for the analysis of human biofluids and tissue

<table>
<thead>
<tr>
<th>Method</th>
<th>Sample type</th>
<th>Description and Use</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC MS</td>
<td>Biofluid analysis</td>
<td>Combination of high-resolution capillary gas chromatography and mass spectrometry. Most commonly used method for detection of metabolic disorders until 1990s when superseded by $^3$HNMR, now superseded by LCMS. Instrumental in providing diagnostic information for numerous inherited diseases Inc. disorders of amino acid metabolism such as phenylketonuria(^{196}), thyroid hormones and organic and fatty acid metabolism(^{197}). Some laboratories worldwide still use GC MS in diagnosis of metabolic diseases.</td>
<td>High resolution. Complex sample preparation, limits on size and type of molecule that can be analysed and slow analysis time.</td>
</tr>
<tr>
<td>LC MS</td>
<td>Biofluid analysis</td>
<td>Liquid chromatography in combination with electrospray ionization mass spectrometry is currently the most common approach towards metabolite profiling in biofluids(^{22}).</td>
<td>Large mass range and high sensitivity. Relatively slow speed of analysis.</td>
</tr>
<tr>
<td>Direct Injection MS</td>
<td>Biofluid analysis</td>
<td>Direct injection or infusion of a sample into the ionization source of the mass spectrometer without prior chromatographic separation. Direct electrospray ionisation combined with tandem mass spectrometry (ESI-MS/MS) is effective and over 35 diseases Inc. newborn screening for PKU are routinely monitored using this technique(^{198}).</td>
<td>Minimal sample preparation. Small sample volume requirement. Process can lead to instrument contamination. Ion suppression may occur and impede meaningful analysis.</td>
</tr>
<tr>
<td>MALDI MSI</td>
<td>Biological tissue sample analysis</td>
<td>Sample covered by a matrix which absorbs light from the subsequent laser irradiation leading to vaporization of the matrix and the sample surface for further analysis(^{199}). Used in clinical laboratories for identification and taxonomic classification of human pathogenic microorganisms more accurately than conventional molecular identification methods(^{200}).</td>
<td>Accurate spatial distribution of analytes possible. Capable of analyzing a wide mass range of large proteins and peptides to small metabolites and lipids (Mass range 100-500 000(^{195})). High spatial resolution (30-75 $\mu$m(^{201})). Sample analysis must occur in instrument vacuum.</td>
</tr>
<tr>
<td>DESI MSI</td>
<td>Biological tissue sample analysis</td>
<td>Sample specific analytes formed by interaction of electrospray</td>
<td>Accurate spatial distribution of analytes</td>
</tr>
</tbody>
</table>
generated charged droplets with the sample surface. Used in research to investigate drug distribution, search for biomarkers of disease and being investigated as a potential tool for histological classification of human tissue sections for medical use.

Possible. Analysis performed in the ambient environment making application to clinical environments easier due to potential for more portable instruments. Matrix not required. Resolution traditionally lower but improving. Large mass range (100 – 66 000 m/z).

**Interpreting Spectra**

A mass spectrum is a plot of ion abundance versus m/z value. The x-axis represents increasing m/z ratio. The y-axis represents the intensity or abundance of each ion relative to the most abundant, which is the tallest peak and set at 100. All other ion peak intensities are shown as a percentage of that most abundant peak.

![](image)

3: Example mass spectrum (m/z range 300-640) demonstrating m/z value on the x-axis and relative abundance of peaks on the y-axis.
2.2 Rapid Evaporative Ionisation Mass Spectrometry

A discussed, mass spectrometry is a useful tool for analysis of tissue lipids; however, in vivo analysis is not possible with current methods, which require tissue preparation. Rapid Evaporative Ionisation Mass Spectrometry is proposed as a tool for in vivo tissue analysis by analysis of the gas phase ionised particles produced at the point of surgical dissection of tissues. The surgical smoke plume is collected from the electrosurgical hand piece, driven by a venturi pump the output of which is orthogonally sampled by the capillary inlet of the mass spectrometer. The capillary introduces the charged aerosol particles into the mass spectrometer. In order to establish an identification system based on mass spectral analysis, large databases comprising histologically assigned data from normal and diseased tissue are constructed by the analysis of ex-vivo human tissue samples prior to surgical intervention. In order to obtain information in real time, the spectra obtained during electrosurgical dissection of tissues are compared to these database entries. REIMS spectra of biological tissues feature a histologically distinct pattern of phospholipids (in the m/z range 600-900), specific to the histological nature of the tissue. The spectral data then undergoes computational analysis ultimately providing information for the surgeon on the biological nature of the tissue being cut.
Figure 4: 1. Overview of REIMS instrumentation and data collection using monopolar electrosurgery as the ionisation method. A/B) Example REIMS mass spectra of human tissue in the phospholipid m/z range (600-900 m/z) with specific phospholipid peaks highlighted. C) Predominant discriminatory membrane phospholipid species detected by REIMS. Image modified from Balog et al Intraoperative Tissue Identification Using Rapid Evaporative Ionization Mass Spectrometry, Sci Transl Med 5, 194ra93 (2013).
REIMS analysis differentiates between tissue types based on differences in lipid profiles, generated by the release of phospholipids on membrane disruption during electrosurgical dissection. In addition to a tool for intra-operative margin detection, these phospholipid profiles, as a measurable end point of de novo lipogenesis and altered lipid metabolism in cancer cells, could provide information on tumour characteristics which could link molecular profiles with biological tumour behaviour.

### 2.3 Electrosurgery

Electrosurgery is the application of a high-frequency alternating electric current to biological tissue as a means to cut or coagulate tissue. It utilizes the heat generated by the passage of a high-frequency alternating electrical current through living tissues. The current is converted to heat by resistance as it passes through the tissue. This allows precise dissection of tissues during surgical procedures whilst limiting blood loss. Principles of electrosurgery are governed by Ohm’s law

\[ \text{Voltage} = \text{Current} \times \text{Resistance} \]

Current is measured as the flow of electrons during a given period of time and Voltage is the force driving the current against the resistance in the circuit. In electrosurgical devices, voltage is provided by the generator, and the current is delivered to the tissues via the tip of the handpiece. Electrosurgery requires a circuit for current to flow and the circuit is completed via the patient return electrode, a low-resistance path through which the current can return to the generator from the patient, reducing the risk of thermal injury due to grounding via alternative routes\(^\text{206}\). Introduction of electrosurgery into medical practice is attributed to William T. Bovie of Harvard University, who developed an innovative electrosurgical unit first used by Harvey Cushing in 1926 during resection of a highly vascular intracranial tumour\(^\text{207}\). Almost 90 years later electrosurgical devices are among the most commonly used tools in modern surgical practice, frequently used in place of a scalpel or scissors for the purpose of tissue dissection. The range of electrosurgical devices available to the modern surgeon is extensive, and includes the most commonly used handheld monopolar device, bipolar forceps, devices for laparoscopic surgery including scissors and dissection hooks, and endoscopic devices such as
polypectomy snares. Monopolar instruments are primarily used for accurate tissue dissection. Energy is precisely delivered to tissues via the tip of the handpiece (active electrode) through the patient and out of the body via the return pad. This allows the electrosurgical instrument to be used in the same manner as a knife, for precise cutting of tissue. Bipolar devices consist of energy delivery forceps containing both the active and return electrode. The target tissue is then grasped between the tips to complete the circuit and energy is delivered to the tissue. Bipolar electrosurgery is commonly used for sealing blood vessels and in cases where unintended dispersal of current to surrounding tissue would lead to significant damage e.g. neurosurgery. The mode of the instrument is altered by modulation of the current output delivered to tissues by the generator. As the output waveforms change, so does the corresponding tissue effect. In general, two modes, cut and coagulation, are used. Cut mode delivers energy as a continuous sinusoidal waveform whilst coagulation mode delivers interrupted current flow exposing the tissue to significantly reduced current therefore requiring higher output voltages. Blend modes offer a degree of each mode. Different modes are used in different surgical procedures and tissue types based upon the surgeon’s preferred effect. Tissue effects that can be achieved with electrosurgery can be divided into 3 groups: cutting, fulguration, and desiccation. Cutting produces the highest current density producing tissue temperatures which rapidly exceed 100 °C which vaporizes the intracellular contents leading to precise division of tissues. Fulguration uses coagulation mode with the tip of the electrode adjacent to the tissues, which generates less heat, creating a coagulum rather than rapidly vaporizing the tissues. The higher voltage coagulates the tissues over a larger surface area. Many surgeons use this mode to cut tissues as it can provide a combination of accurate tissue dissection and coagulation of vessels to reduce blood loss. Desiccation occurs when coagulation mode is applied in direct contact with the tissue surface. Contact with tissues leads to reduced current concentration and heat production leading to formation of a coagulum. A variety of effects can be produced by each mode depending on how the surgeon chooses to manipulate the interaction of the tissues with the electrode. A byproduct of electrosurgical dissection or coagulation is surgical smoke, produced when the heat from the device heats target cells to the point of membrane rupture. This smoke plume is released into the general atmosphere of the operating theatre and eventually cleared
by the operating theatre ventilation and frequent whole room volume air changes. If offensive or distracting, the smoke may be extracted at the time of dissection either by a specialised suction device attached to the electrosurgery handpiece or by close application of the standard operative suction device. The smoke plume contains tissue specific charged particles, and these are harnessed and used for tissue discrimination during REIMS analysis.

### 2.4 Desorption Electrospray Ionisation Mass Spectrometry (DESI)

As described, mass spectrometry imaging (MSI) is being increasingly applied to investigate the spatial distribution of biological molecules in tissue sections. DESI MSI can be used to spatially map ionic species, including small metabolites and complex lipids, in unprepared tissue sections in the ambient environment. During DESI analysis, ionisation takes place on the surface of the tissue sample. Ionisation of sample analytes is achieved by directing a spray of charged solvent droplets (electrospray) onto the surface of the tissue. On impact these droplets take up molecules from the sample surface, after impact the solvent droplets revert back to the gas phase leading to ionization of the dissolved surface molecules, which are transported into the mass spectrometer, via the inlet capillary, for analysis. Each tissue sample is analysed in a linear (raster pattern) fashion so that an image of the sample, based solely on the chemical species measured, can be reconstructed\(^{202}\). This method allows the same, rather than adjacent, tissue sections to form the basis for the histological and mass spectral image. Analysis of MSI data requires precise co-registration of features from MS and optical images. An in house image alignment algorithm\(^{208}\) creates a precise pixel-to-pixel match for the 2 images. Once the images are co-aligned, individual tissue types within the histological image are selected, and there is automatic co-selection of the corresponding chemical co-ordinates on the MSI image. Using this method, areas of interest such as cancer cells, stroma and histologically normal epithelial cells can be selected and corresponding mass spectral profiles can be examined. Overall chemical profiles of the lipid composition in each region can also be generated by combining and averaging the spectra from each region prior to analysis. The DESI data generated by spatially resolved analysis of each tissue sample forms a replicate image of the histological tissue section, based on inherent differences in the mass...
spectral profiles of individual tissue types within the sample. The spectral data within each DESI image, with its co-registered histological classification can then be subject to multivariate analysis.

Figure 5: Schematic diagram of DESI instrumental setup with mobile electrospray source, inlet capillary and sample stage. b) Ionisation of the tissue sample surface occurs in a raster pattern. Each movement of the sample under the sprayer generates its own, topographically accurate, mass spectrum, representing the chemical composition of the sample surface in that area. Using this information, a chemical image of the sample is formed, based on the MS profile derived from each sampling point. Figure modified with permission from Mass spectrometry sampling under ambient conditions with desorption electrospray ionisation. Takats et al. Science. 2004 Oct 15;306(5695):471-3

2.5 Data Analysis

2.5.1 Mass Spectral Data Interpretation

REIMS analysis generally produces singly charged, deprotonated (M-H) molecular ions which are measured in negative ion mode. As such, the recorded signal for any detected molecular species corresponds to its molecular mass, which simplifies data interpretation. The separation and detection of charged species requires between 150 and 500 ms per iteration so each mass spectrum is generated within 0.5s of tissue sampling. Ions are predominantly detected in the 600 to 900m/z range, corresponding to 600 to 900 Da, if all are in the expected singly charged state. Data analysis can be divided into training and classification phases. In the training phase, the tissue classification algorithm is trained to extract molecular ion patterns of histologically classified tissue types, and these are subsequently used for prediction of unknown tissue types in the classification phase.
2.5.2 Database Construction

REIMS spectra of biological tissues feature histologically distinct patterns of phospholipids, specific to each tissue type. In order to identify tissue in near real time, tissue specific spectral information gathered during *in vivo* electrosurgical dissection is compared to previously constructed, histologically assigned spectral databases of healthy and diseased tissue types. In order to build tissue specific spectral databases, *ex vivo* tissue samples are sampled using the monopolar electrosurgery REIMS setup described above, and the spectral data stored. The tissue samples are then placed in formalin and undergo sectioning and H&E staining. The H&E stained slides, with the electrosurgery cut visible, are examined by a histopathologist who provides a tissue diagnosis and each spectral datafile is assigned a tissue type. More detailed clinical characteristics are obtained from the formal histology report for each resected specimen.

Figure 6: Example REIMS spectral data files obtained from *ex vivo* sampling of colorectal adenocarcinoma and normal adjacent mucosa with the phospholipid region of interest (600-900 m/z) highlighted. Magnified views of the corresponding H&E stained slide of each tissue sample are shown, with visible electrosurgery cuts, used for histological classification of the spectral data file within the database.

2.5.3 Multivariate Analysis

Prior to multivariate analysis, the raw REIMS data is pre-processed for optimized information recovery.
This is necessary because the data generated from profiling experiments contains a degree of unwanted background signal or “noise” which provides no additional information. Additionally very small day to day variation in instrument conditions can lead a degree of mass drift - poor alignment of peaks, generated from the same molecule when sampled at different times. The raw spectral data is imported into MATLAB (R2014a) into a purpose built toolbox, designed for interrogation of clinical mass spectral datasets. The compiled spectra are normalized to their total ion count to reduce the impact of variation in overall signal intensity unrelated to spectral pattern. Alignment is carried out to offset the effect of mass drift; where peaks naturally drift between sample runs due to sample build up within the instrument and minute day to day instrumental variation. The data is log transformed to account for large variability differences between small and large peaks across mass spectra. The dataset is then subjected to dimension reduction via Principal Components Analysis, an unsupervised variable reduction method, which reduces the number of observed variables into a smaller number of orthogonal principal components, which account for the majority of the variance in the dataset. This enables visualisation of classes within the biological data set based on the inherent similarity/dissimilarity of samples with respect to their biochemical composition, by visualisation of clustering behaviour. Linear discriminant analysis is then used to maximize the separation among different data groups by the identification of orientations responsible for maximal separation of each class. While PCA is an unsupervised algorithm not taking the histological assignment of samples into account, LDA is a supervised approach, which utilizes this extra descriptive information when modeling the dataset. The objective of LDA is to transform the multidimensional data space in a way that individual groups of data, in this case corresponding to distinct histological classes, are tightly clustered with maximized separation among classes. The resulting LDA space is used for the classification of spectra from unknown tissue samples by localizing the corresponding data point in the data space, and calculating its distance from the centre of histologically assigned data groups. Mahalanobis distance measurements are calculated from the unknown data point to the centers of each of the histologically assigned data clouds. The data point is classified into the closest group using a certain threshold value, which depends on the average distance of database entries from the centre. Since all but the distance
calculation is carried out prior to analysis of an unknown sample, subsequent identification of unknown tissue types can be achieved within a few seconds.

The workflow for the analysis of pre-processed REIMS data in demonstrated in Figure 7.

2.5.4 Interpretation of Results

For improved interpretation and visualisation, results of REIMS tissue analysis throughout this work are presented as confusion matrices and receiver operating characteristic (ROC) curves. Results are displayed in confusion matrices to allow visualisation of the diagnostic performance of each supervised analysis. Each column of the matrix represents sample datapoints in a predicted class, while each row represents sample datapoints in an actual (target) class. This allows visualization of the number of correctly and incorrectly classified samples and therefore the number of true and false positives and true and false negatives. An overall classification accuracy is also provided which represents True positives + True negatives/Total number of samples, or the percentage of samples correctly assigned.
The boxes appear in various shades of black, grey and white to provide a visual representation of the accuracy of the test. For example, if all samples in group “1” are correctly classified as such and there are no false negatives then the true positives box will appear black and the false negatives box will appear white. If 50% are misclassified then both boxes will appear the same shade of grey.

ROC curves are also used to demonstrate the accuracy of each supervised analysis. ROC curves illustrate the performance of a binary classification system, such as a diagnostic test to identify cancer. The true positive rate, or sensitivity is plotted, along the y-axis, against the false positive rate, or 1-specificity plotted along the x-axis. The most accurate test would yield a point at coordinate 0,1, which would represent 100% accuracy (no false positives and no false negatives.) Random chance yields a
diagonal line through point 0.5, 0.5. Therefore, an AUC of 0.5 represents a result no better than random chance and the closer to 1, the more accurate the test.

![ROC curve, AUC = 0.94796](image)

Figure 9: Example ROC curve with high AUC representing an accurate discriminatory test

### 2.6 Device Development and Regulation within the UK

CE marking of a medical device ensures that the product meets the essential requirements of all relevant European Medical Device Directives and is a legal requirement to place a device on the market in the European Union. The Medical Device Directives were harmonized in the EU in 1990’s and relate to the safety and performance of medical devices. The three directives include:

- Directive 90/385/EEC regarding active implantable medical devices (AIMD)
- Directive 93/42/EEC regarding medical devices (MDD)
- Directive 98/79/EC regarding in vitro diagnostic medical devices (IVDD)

A Medical Device (MD) is defined as

“Any instrument, apparatus, appliance, material or other article, whether used alone or in combination, including the software necessary for the proper application, intended by the manufacturer to be used for human beings for the purpose of

- diagnosis, prevention, monitoring, treatment or alleviation of disease,
- diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap,
• investigation, replacement or modification of the anatomy or of a physiological process,

• control of conception

and which does not achieve its principal intended action in or on the human body by pharmacological, immunological or metabolic means, but which may be assisted by such means. Conversely, an In Vitro Diagnostic Medical Device (IVDD) encompasses any device including reagents, products, calibrators, control materials, kit, instruments, apparatus, equipment, or systems, intended to be used, in vitro, for the examination of specimens derived from the human body, including blood and tissue donations, for the purpose of providing diagnostic or prognostic information.

Medical devices may be classified as Class I, Is, Im Class Ila, IIB and III, with Class I devices deemed to be the lowest risk requiring the least regulatory control and Class III covering the highest risk products. The assigned classification determines the level of assessment that is required for approval and CE marking. All apart from Class I (excluding Is, products marketed as sterile or Im, those with a measuring function) require the intervention of third party, or Notified Body. The class of a device is determined by multiple factors including the invasive nature of the device, the duration of use, whether the device is an energy source and the part of the body affected by the device - those in contact with the central nervous system are automatically placed in category III.

Based on the definitions outlined above, both the modified electrosurgery handpiece and the REIMS capable mass spectrometer, with its built in analytical software, would be classified as medical devices should they be used for in vivo diagnostic purposes. The Medicines and Healthcare products Regulatory Agency (MHRA) are the competent authority for registration of medical devices within the UK. As part of the process of developing a novel medical device, I contacted the MHRA for advice around CE marking and regulatory approval during the developmental and experimental stage of REIMS development. They advised that devices which are manufactured by healthcare establishments and are only used on their own patients are exempt from the requirements of the medical devices regulations, as are devices which are used solely for the purpose of research. Guidance (Bulletin 18 - The Medical
Devices Regulations: Implications on healthcare and other related establishments) produced by the MHRA states that a CE mark should not be placed on a device if it is

• a custom made device, although it must still meet the requirements of the directives
• a product undergoing clinical investigation. These products must be labelled clearly as exclusively for clinical investigation and meet the requirements as far as possible. When using such a device all precautions must be taken to protect the health and safety of patients
• an in vitro diagnostic medical device for performance evaluation
• a non compliant device may be used in exceptional circumstances e.g. humanitarian grounds.

During the manufacturing process, the Xevo G2-S iknife QToF instrument, was built to comply with en 5501 - industrial, scientific and medical (ISM) equipment, classification Group 1 Class A. The classification was assigned in accordance with IEC CISPR 11 Industrial Scientific and Medical (ISM) instruments requirements. Group 1 includes products that intentionally generate and/or use conductively coupled radio-frequency energy necessary for the internal functioning of the equipment. Class A products are suitable for use in commercial locations and can be directly connected to a low-voltage, power-supply network

At the present time, the mass spectrometer and associated ancillary equipment are for research use only. Although the instrument will be used in the operating theatre environment, this will be for the purpose of collecting data for research and any accrued data will not be used to provide a diagnosis. As such the equipment has not been CE marked according to EU medical device directive, however, has been clearly labeled as a tool for research purposes only and complies with ISM requirements. Application for CE marking would be required if the equipment were to be marketed as a medical device in the future.
3. Hypothesis, Aims and Objectives and Study Design

3.1 Hypothesis

The overall hypothesis of this body of work is that the structural phospholipid content of mammalian tissues shows both histological and pathological specificity; direct (or ‘ambient’) ionisation of these phospholipids by mass spectrometric methods allows for unambiguous identification of tissues and pathologies in near real time, and provides phenotypic data of importance in the surgical management of cancer.

3.2 Aims

1. To develop, optimize and validate the instrumentation and analytical methodology for the analysis of human tissues by Rapid Evaporative Ionisation Mass Spectrometry, suitable for future use in the operating theatre environment.

2. To build a histologically validated spectral database for the identification of benign and malignant breast and colorectal tissue using fresh ex vivo tissue, and validate this using clinical and molecular platforms (histology and DESI imaging).

3. To use this data to develop a standardised methodology of data handling for REIMS spectral and clinical data for future REIMS studies.

4. To validate the diagnostic performance of the REIMS-based iKnife system as a tool for the ex vivo differentiation of malignant and non-malignant breast and colorectal tissue.

5. To determine whether REIMS analysis of malignant cell membrane structural phospholipid changes can provide prognostic information, based on validated clinical prognostic indicators such as tumour grade and lymphatic involvement.

6. To perform a feasibility pilot study into the translation of REIMS technology to an endoscopic tool for the intraluminal identification of colorectal dysplasia and malignancy.
3.3 Study Design

The study design was optimized to first develop the electrosurgical device, MS interface and MS instrumentation required for REIMS analysis of human tissues. This instrumental setup was then validated in the building and prospective testing of ex vivo REIMS databases of colorectal and breast tissue. The REIMS platform developed during initial instrument development section was not suitable for the analysis of human breast tissue, therefore, further instrument experimental work in instrument design and development was carried out to develop a system which was compatible with the analysis of breast tissue. This analytical platform was then validated by the building and prospective testing of an ex vivo REIMS breast cancer database. Both DESI and REIMS analysis was performed on a small subset of colorectal tissue to determine whether similar chemical species were detected in the analysis of the same tissue type using different platforms and a method was developed for multi-platform utilization of data generated from MS analysis of human tissues. Finally, the instrument and analytical strategy developed for the analysis of colorectal tissue was applied to a tool for endoscopic analysis of colorectal tissues and the system was tested by the building of a small database to determine feasibility of translation of REIMS technology into an endoscopic device.

3.3.1 REIMS Instrument Development

The primary objective of this experiment was the development of an electrosurgery-mass spectrometer based system suitable for in vivo REIMS analysis of human tissue within operating theatres in the UK. This was a feasibility study which utilized human colorectal tissue samples for development and testing of the system. Validation was subsequently performed in later experiments as part of a prospective pilot analysis of the diagnostic power of ex-vivo REIMS sampling of human colorectal and breast cancer tissue.

Design Overview

1. An iterative design of a REIMS capable modified electrosurgical handpiece for the efficient transfer of a surgical smoke plume to the mass spectrometer, and equivalent functionality of a commercially available handpiece.
2. The creation of a mass spectrometry interface, for the provision of spectral data of sufficient quality for the discrimination of histologically distinct human tissue types based on prospective optimisation against the ex-vivo analysis of human colorectal and breast tissue.

3. Qualitative analysis of current criteria for the development of a mass spectrometer which conforms to safety guidelines for operating theatre equipment and which is suitable for use within UK NHS operating theatres in terms of its practical size and mobility. This included a compatibility analysis of power and gas supplies and the prospective modification of a commercially available instrument in collaboration with the manufacturer.

The primary outcome measures were:

1. Clinical suitability and regulatory standardisation
2. Machine stability for prospective use in a clinical environment
3. Spectral data quality based on noise to signal ratio and reproducibility of identified molecular species.

This prospective, feasibility study was not powered to meet clinical diagnostic endpoints since it was a qualitative analysis. Work was carried out in collaboration with Waters Corporation and their engineers to ensure the highest commercial quality was achieved.

**3.3.2 REIMS Analysis of Colorectal Cancer**

REIMS analysis was applied to colorectal tissue samples to determine whether the lipid profile of human colonic mucosa can be measured by REIMS, and to ascertain its diagnostic accuracy for primary adenocarcinoma. A secondary analysis was performed to determine if REIMS lipidomic is able to differentiate adenocarcinoma, dysplasia and healthy mucosa and whether the lipid signatures detected provide molecular phenotyping information of prognostic use using standardized histological biomarkers as the gold standard.

**Design Overview**

This was a prospective observational study performed in a single centre at St. Mary’s Hospital, Imperial College London. The primary outcome measure was the diagnostic accuracy of REIMS in the ex vivo
identification of colorectal cancer and normal adjacent mucosa based on analysis of lipid profiling data. Secondary outcome measures were the diagnostic accuracy of REIMS in the ex vivo identification of dysplasia and distinction from invasive cancer and normal adjacent mucosa and REIMS differentiation of tumour tissue based upon T stage, N stage, KRAS mutation and EMVI status. This analysis was not powered as the diagnostic accuracy of the technique was not fully determined.

An experimental protocol for patient recruitment, tissue collection and sampling of colorectal and breast tissue that did not compromise clinical care and produced spectral data of sufficient quality for subsequent analysis was produced and distributed. Spectral databases were built using REIMS analysis of histologically validated ex vivo human tissue samples of colorectal adenocarcinoma and normal adjacent mucosa. Data analysis was performed using PCA, LDA, Mahalanobis distances and leave one patient out cross validation. Identification of lipid species responsible for the differentiation of colorectal adenocarcinoma and normal adjacent mucosa was performed using ANOVA of cross validation rounds with a p value of 0.05. Histological validation was carried out by an independent pathologist who was blinded to REIMS classification of tissue samples. Clinicopathological information was collected, with ethical approval, from patients’ hospital records and formal histopathology reports.

3.3.3 REIMS analysis of Breast Cancer

REIMS analysis was applied to samples of human breast tissue to define the REIMS lipid profile of breast cancer and healthy breast tissue and to determine whether REIMS lipid profiling of breast tissue provides tissue specific information sufficient for the ex vivo identification of invasive breast cancer. No previous validation work had been performed in the use of REIMS analysis on breast tissue and there was no available data on the effect of the high lipid content of human breast tissue on the quality of the spectral data obtained. REIMS analysis of breast tissue varies considerably from that of other tissue types such as the colonic mucosa. Healthy human breast tissue is composed predominantly of fat and the internally charged aqueous droplets formed during electrosurgical dissection of other tissue types are not seen during analysis of breast tissue, instead large lipid droplets are formed. These droplets do not dissociate into smaller charged particles at the collision surface of the mass spectrometer in the
same way, and REIMS analysis of lipid rich healthy breast tissue therefore yields no MS signal. Prior to performing any experimental work to determine the accuracy of REIMS in the identification of histologically distinct breast tissue, it was necessary to first develop instrumentation and analysis methodology that produced high quality spectral data during REIMS analysis of the human breast.

**Design Overview**

The study was performed in two parts:

1. Creation of a stable REIMS instrumental platform suitable for the analysis of lipid rich human breast tissue
2. A prospective, single centre observational study of REIMS analysis of breast tissue.

The primary outcome measure was the ex vivo tissue discriminatory capacity of REIMS in the identification of invasive breast cancer. Secondary outcome measures were the diagnostic accuracy of REIMS in the ex vivo identification of benign breast lesions and detection of an association between lipid profiles of human breast cancer, as measured by REIMS, and tumour characteristics of prognostic significance including tumour type and histological grade, oestrogen receptor status and HER2 overexpression. A novel atmospheric interface with an isopropanol (IPA) infusion combined with the venturi pump was developed, which produced spectral profiles with phospholipid signal from the analysis of both healthy and malignant breast tissue. Spectral databases were built using REIMS analysis of histologically validated ex vivo human tissue samples of invasive breast cancer, benign breast lesions and healthy breast tissue. Data analysis was performed using PCA, LDA, Mahalanobis distances and leave one patient out cross validation. Identification of lipid species responsible for the differentiation of breast cancer and normal adjacent mucosa was performed using ANOVA of cross validation rounds with a p value of 0.05. Histological validation was carried out by an independent pathologist who was blinded to REIMS classification of tissue samples. Equivocal samples were examined by a second histopathologist. Clinicopathological information was collected, with ethical approval, from patients’ hospital records and formal histopathology reports.
3.3.4 DESI REIMS Data Cross Validation

DESI MSI provides high volume, highly spatially resolved (40-75 µm), data from the ionization and analysis of lipid species in human tissue samples, but is not suitable for \textit{in vivo} use. REIMS also relies upon the detection of ionized lipid species, produced during electrosurgical dissection however REIMS suffers from poorer spatial resolution, inefficient utilization of tissue samples and is destructive, therefore the exact location sampled by REIMS cannot undergo histological examination whereas DESI MSI analysis leaves behind intact tissue morphology for histological examination. For these reasons DESI is a superior tool for database construction, however, REIMS has potential as a clinical tool for near real time analysis. It is hypothesized that, despite different ionization mechanisms and hence differences in the mass spectral profiles produced, DESI MSI detects similar chemical species to REIMS in the analysis of equivalent tissue samples. The aim of this experiment was to determine whether high resolution, spatially resolved spectral databases built using DESI data, can be used in the prospective classification of tissue sampled by REIMS. The major challenge in achieving this lies in the differences between ionization methods and instrumental platforms, which introduce platform specific variation to the datasets. This required the development of a cross platform data fusion method to obviate individual platform effects and allow direct comparison of data analysed by both methods. This algorithm was applied to DESI and REIMS analysis of colorectal tissue to determine whether tissue classification is possible across multiple platforms.

Design Overview.

A spatially resolved DESI MSI database of human colorectal tissue was built by DESI analysis of samples of colorectal adenocarcinoma and normal adjacent mucosa and the discriminatory capacity of this database was tested. Lipid species responsible for the DESI differentiation of colorectal adenocarcinoma and normal mucosa were identified and compared to significant peaks seen in REIMS analysis of corresponding tissue. A data fusion method was developed to reduce MS platform-specific variation sufficiently enough for the accurate identification of colorectal adenocarcinoma and healthy mucosa sampled with REIMS using pre-built DESI MSI datasets.
3.3.5 iEndoscope

REIMS technology was translated to an endoscopic tool to determine whether a REIMS endoscopic snare system is suitable for transfer of the ionic gaseous product of colorectal electrosurgical dissection into a mass spectrometer for REIMS analysis to take place.

Design Overview.

This was a feasibility pilot study performed in a single centre. The primary outcome measure was the development of an endoscopic snare system which produced spectral data of sufficient quality for the identification of colorectal adenocarcinoma from the *ex vivo* REIMS analysis of colorectal adenocarcinoma and normal mucosa. Secondary outcome measure was the REIMS identification of colonic wall layers from *ex vivo* analysis of samples of full thickness healthy colon. Spectral databases were built using REIMS analysis of histologically validated *ex vivo* human tissue samples of colorectal adenocarcinoma and normal adjacent mucosa and independently sampled colonic wall layers. Data analysis was performed using PCA, LDA, Mahalanobis distances and leave one patient out cross validation. Histological validation was carried out by an independent pathologist who was blinded to REIMS classification of tissue samples.
4. Materials and Methods

This study was subject to research governance structures provided internally within the department of Biosurgery and Surgical technology in conjunction with the section of Biomolecular medicine for the safe use of samples and machinery.

4.1 Patient Recruitment

Colorectal and breast tissue samples utilised in the following studies were obtained from patients undergoing resectional surgery within Imperial College Healthcare NHS Trust. Ethical approval was obtained from the South East London National Research Ethics Committee (Study ID 11/LO/1686). Patients undergoing surgery for breast and colorectal disease were prospectively recruited prior to surgical intervention. Patients were supplied with verbal and written information about the study and subsequently provided written, informed consent. Prior to beginning the study an experimental protocol for the collection, REIMS analysis and histological examination of colorectal and breast tissue was written, distributed and agreed with the surgeons and histopathologists involved in the care of the patients. All patients over the age of 18 years undergoing surgical resection for benign or malignant colorectal disease who were able and willing to provide informed consent were eligible for inclusion into the study. Patients who had undergone neoadjuvant therapy were included, and this was documented in the corresponding clinical dataset in addition to gender, age, medical comorbidities, drug history and primary pathology including tumour location, histological subtype, grade, stage, presence of EMVI and gene mutations. Male and female patients over 18 years of age undergoing breast surgery for benign, pre-malignant and malignant disease who were able and willing to provide informed consent were eligible for inclusion into the study. Patients who had undergone neoadjuvant therapy were included, and this was documented in the corresponding clinical dataset, as was: medical history, drug history, menopausal state, ethnicity, smoking and alcohol status, lifetime use of hormone therapy and primary pathology including tumour location, histological subtype, grade, stage, hormone receptor and HER2 status.
4.2 Tissue Bank Sampling

Tissue samples were also sourced from the Imperial College NHS Trust Tissue bank. Imperial College Healthcare Tissue Bank was first established in 1998 and collects, stores and distributes tissue collected, with written consent, from patients undergoing surgical procedures throughout the Trust. 20 matched pairs of colorectal cancer tissue and normal adjacent mucosa were sourced for use in instrument and analysis methodology development. 40 samples of breast tissue (10 normal, 10 benign, 10 invasive cancer, 10 in situ malignancy) were also sourced for use in REIMS instrumentation and methodology development for analysis of breast tissue. 15 tissue samples (12 invasive cancer, 2 fibroadenoma and 1 normal glandular breast tissue) were included in the final breast cancer REIMS dataset. Samples sourced from the tissue bank were snap frozen at the time of harvest and stored at -80°C. This tissue was sourced under ethical approval (11/LO/1686).

4.3 Tissue Sample Collection, Storage and Analysis

Resected surgical specimens were collected from the operating theatre and transported immediately to the histopathology department where appropriate fresh tissue samples were retrieved by a histopathologist. Where possible, samples of tumour or adenoma/benign lesions and normal adjacent mucosa (NAM) >5cm from tumour edge were sampled for each patient using cold steel dissection. After collection, tissue samples were immediately frozen at -80°C. Prior to REIMS analysis, samples were thawed at room temperature. The type of tissue available for research sampling was heavily dependent on the size of the tumour and extent of normal tissue contained within the surgical specimen. For example, following neoadjuvant therapy, there might have been insufficient tumour tissue to safely provide a sample for research. Similarly, following a wide local excision for breast cancer, the extent of normal breast tissue contained within the sample might have been too small to safely provide a sample. Therefore, it was not possible to consistently obtain samples of both tumour and healthy tissue from each patient.
4.4 Clinical Phenotyping Laboratory

A custom built Waters Xevo G2-S QTof i-Knife mass spectrometer was installed in a purpose built clinical phenotyping laboratory within the operating theatre suite in St Mary’s Hospital, London. The laboratory also housed 2 $^1$H NMR spectrometers for biofluid analysis. Here, within the operating theatre environment, the instrument could be tested for both medical functionality/compatibility with the operating theatre environment and ion collection/transfer efficiency. Elements tested included, efficiency of splitting the main medical air supply to anaesthetic machine to drive the venturi pump, locating appropriate space within the operating theatre to house the instrument that did not compromise functionality and access to power supply of other surgical instruments even if not in use, theatre staff tolerance of noise generated by the instrument and impact of venting the vacuum and moving the instrument on data quality.

4.5 REIMS Tissue Sampling

4.5.1 Colorectal Tissue

Each surgical specimen in its entirety was collected fresh, without formalin, from the operating theatre and transported to the histopathology lab at St Mary’s Hospital. Samples of colorectal adenocarcinoma, adenoma and adjacent mucosa more than 5cm from the tumour were excised from the surgical resection specimen by a histopathologist to ensure post-operative histopathological processing, staging, and reporting of the specimen, in accordance with hospital Standard Operating Procedures and Royal College of Pathologists datasets, was not compromised. Tissue was divided so that it could be embedded in standard-size paraffin histology blocks. Tissue specific spectral data were acquired by sampling each tissue specimen with a modified handheld monopolar electrosurgical handpiece connected to a commercially available electrosurgery generator (Valleylab SurgiStat™ II) using cut mode, power 30W. The hand piece was 155mm in length with a maximum circumference of 50mm with a standard cutting tip 35mm in length x 2mm in diameter. The blade was embedded in an additional 15mm long piece of stainless steel tubing with a 2mm internal diameter. The stainless steel tubing was
connected to a central aspiration line, incorporated into the handpiece design, to allow aspiration of the smoke plume. The aerosol formed by electrosurgical dissection was transferred to the mass spectrometer via a 2 m long section of PTFE tubing with 3.2mm outer and 1.5mm inner diameter. The mass spectrometer was equipped with a modified atmospheric interface featuring heated capillary atmospheric inlet and orthogonal Venturi-pump aerosol transfer device. The aerosol is transferred using the Venturi air jet pump operated at a nitrogen inlet pressure of 5 bars in order to establish sufficient suction force for aerosol transport. Exhaust of the Venturi pump is directed orthogonally to the inlet of the mass spectrometer. The inlet capillary introduces the charged aerosol particles into the instrument (Xevo G2-S i-Knife QToF) where they undergo further dissociation at the atmospheric interface, which consists of a coil of Kanthal, an iron-chromium-aluminium alloy able to withstand high temperatures and used frequency in heating elements, heated to 800°C, prior to entry into the instrument for analysis. Mass spectra were recorded in negative ion mode, in the m/z range 200-1000. The instrument was controlled via an embedded computer system, which controls all functions and also displays total ion count and raw mass spectra during analysis of tissues. The touch screen of the embedded computer system will also act as a display monitor during future in vivo analysis. Power supply is from building main electricity supply (240v). The exhaust cooling air is filtered through an integrated HEPA filter.

4.5.2 Breast Tissue

Each surgical specimen in its entirety was collected fresh, without formalin, from the operating theatre and transported to the histopathology lab at Charing Cross Hospital. Samples of invasive breast cancer and normal healthy breast tissue were excised from surgical resection specimens by a histopathologist as above. Where possible, healthy breast tissue was sampled more than 1cm from the tumour edge, however, in cases of breast conserving surgery this was not always possible due to the small volume of normal breast tissue within the specimen. Tissue was divided so that it could be embedded in standard-size paraffin histology blocks. REIMS data was collected using a needlepoint diathermy tip (35mm in length x 0.5mm in diameter) to allow accurate sampling and precise histological evaluation of the tissue.
post analysis. In small tissue samples, one needlepoint was sampled, in the case of larger tissue samples, several needlepoints were analysed (separate spectra for each point obtained) and the tissue was inked at the right lateral margin to allow correct orientation of the corresponding microscope slide (Figure 10).

![Image](image1.png)

**Figure 10:** Sample of breast tissue demonstrating 4 distinct needle electrode sampling points and the corresponding microscopic image of normal fibro-fatty breast tissue with a visible electrosurgery sampling point used to histologically classify the MS datafile within the database.

Tissue specific spectral data were acquired by sampling each tissue specimen using the monopolar handpiece setup as described above. Ionic species were transferred to the atmospheric interface of the mass spectrometer as described above. In order to generated ionized particles from the large lipid droplets generated by REIMS analysis of breast tissue, continuous isopropanol infusion at a rate of 450 microlitre/hr was added to the nitrogen venturi flow (discussed in Chapter 5.1). The venturi exhaust, rather than being positioned orthogonally to the inlet capillary of the mass spectrometer, was transported directly into the inlet capillary via a piece of PTFE tubing 10mm in length, 3.2mm outer and 1.5mm inner diameter. Charged aerosol particles introduced into the instrument undergo further dissociation at the heated coil atmospheric interface prior to analysis. Mass spectra were recorded in negative ion mode, in the m/z range 200-1000.

### 4.5.3 iEndoscope Tissue Sampling

A commercially available disposable electrosurgical polypectomy snare (SD-210U-15; Olympus Medical Systems Corp) with a 15mm loop diameter was modified for use as the electrosurgical dissection tool. Electrosurgical power was obtained from a SurgiStat™ II Electrosurgical Generator. The snare was
adapted to allow transfer of gaseous content from the dissection site by placement of a BD Connecta™ three-way stopcock (Becton Dickinson UK Ltd) approximately 20 centimetres distal to the snare handle and creating puncture holes in the plastic insulating layer of the snare, adjacent to the tip to allow improved aspiration of the aerosol. The snare was then coupled to the mass spectrometer (Xevo G2-S i-knife) via a 1m long, 3.2mm outer and 1.5mm inner diameter section of PTFE tubing placed into the third arm of the stopcock. Instrument venturi setup and aerosol transfer was as per colorectal tissue sampling.

![Diagram](image.png)

11: a) Schematic diagram of the electrosurgical snare system within the colonoscope, connected to the mass spectrometer for smoke plume analysis. b) Modified endoscopic snare with puncture holes in the plastic insulator and the addition of PTFE tubing via a three-way stopcock valve distal to the snare handle.

Samples of colorectal adenocarcinoma and adjacent mucosa more than 5cm from the tumour were excised from the surgical resection specimen. For analysis of colonic wall layers, a further 2 x 2cm sample of full thickness normal colon was excised more than 5cm from the tumour edge.

Spectral data was acquired by sampling each *ex vivo* tissue specimen with the modified endoscopic snare passed down the instrument channel of the colonoscope. Power supply to the snare was controlled by a footpedal and tissue was sampled on cut mode with a power of 20W. During analysis of
individual colonic wall layers, data was consecutively acquired from each anatomical layer within a single sample of normal colon by injecting distilled water beneath the mucosa using a 5ml syringe and 25G needle to lift and isolate it from the layers beneath. Each layer was subsequently sampled by making an electrosurgical incision confined to that layer. Samples of adenocarcinoma and normal adjacent mucosa were sampled individually with the modified endoscopic snare with settings as described above.

4.6 Tissue Processing and Histological Examination

After analysis, tissue samples were fixed in 10% neutral buffered formalin for 7 days. Fixed samples, including the electrosurgical burn point, were cut into appropriate portions and placed in embedding cassettes. Each sample was dehydrated through a series of graded ethanol baths to displace the water, and then infiltrated with paraffin wax. The infiltrated tissues were then embedded into wax blocks, 5 μm thick paraffin sections were cut using a microtome, fixed onto glass slides and H&E stained. Each slide was examined by a histopathologist for validation proposes.

4.7 REIMS Data analysis, Database Construction and Multivariate Analysis

The workflow for the construction of a REIMS spectral database is illustrated in Figure 7. REIMS raw spectral data were converted to imzML format (MSConvert) and imported into Matlab (R2014a) for processing within a toolbox designed in-house (Dr Kirill Veselkov). The compiled spectra were normalized to their total ion count to reduce the impact of variation in overall signal intensity unrelated to spectral pattern and then log transformed to account for large variability differences between small and large peaks across mass spectra. Discriminating models were then subjected to dimensionality reduction via Principal Components Analysis to reduce the number of observed variables into a smaller number of orthogonal principal components, which account for the majority of the variance in the dataset. The first 30 principal components were kept for further analysis. Linear discriminant analysis was then used to maximize the separation between different classes and the resulting LDA space was used for classification of spectra from unknown tissue samples using Mahalanobis distance calculations.
Leave-one-patient-out cross-validation was performed. All mass spectral data belonging to one patient were left out of the sample set and a new model was calculated using the remaining data. The withheld data was projected into the new model and classified as one tissue type using Mahalanobis distances calculations, performed between the unknown sample point and calculated class centres. This process was repeated for each individual patient. To identify statistically significant peaks responsible for prediction into each class, for each cross-validation iteration, analysis of variance (ANOVA) was employed to only select the features most statistically significant (p value < 0.05) for PCA-LDA classification into each group. ANOVA calculates the mean spectra for each group and then compares the between-group-variations to within-group-variation. Those features with a strong difference between these two variations are considered statistically significant. Exact m/z values were retrieved from the raw data and these values were used for compound identification via database search (Metlin).

4.8 DESI MSI Analysis

4.8.1 Tissue Sampling

Tissue specimens were retrieved during surgical procedures and immediately frozen at -80°C. Prior to analysis, frozen tissue sections (thickness: 15 µm) were prepared from each sample using a cryostat. Tissue samples were stored in the cryostat, at a temperature of -20°C throughout the process and attached to stainless steel chucks using distilled water as a freeze mounting medium. Each section was placed, uncovered, on a glass slide and brought to room temperature prior to DESI analysis.

4.8.2 DESI Instrumentation

A home-built DESI ion source was coupled to a high resolution orbital trapping mass spectrometer (Exactive, Thermo Scientific GmbH, Bremen, Germany) The ion source was operated at a spatial resolution of 100 µm using nitrogen as nebulising gas at pressure of 4 bar, high voltage of 4.5 kV, a solvent mixture of methanol and water in a ratio 95:5 and a flow rate of 1.5 µl/min. The height distance between the DESI sprayer and the sample surface was set to 4mm, and the distance between the
sprayer and the inlet capillary set at 14mm. Mass spectra were recorded in negative ion mode, in the m/z range 200-1000. After measurement, tissue sections analysed by DESI were H&E stained, digitally scanned at high resolution using a NanoZoomer 2.0-HT digital slide scanner and the images were annotated by a histopathologist.

4.8.3 DESI MSI/Histological Data Co-Registration

Prior to analysis, co-registration of features from each MS and optical image was carried out using an in-house image alignment algorithm\textsuperscript{208}. Once co-alignment had been performed, individual tissue types within the histological image were identified and annotated by a histopathologist, and there was automatic co-selection of the corresponding chemical co-ordinates on the MSI image. The spectral data within each DESI image, with it’s co-registered histological classification was then uploaded into a custom built toolbox\textsuperscript{208} operated using the Matlab 2014a processing environment for further analysis.

4.8.4 DESI Data Analysis, Database Construction and Multivariate Analysis

Raw mass spectrometric data were converted to imzML format (MSConvert) and imported into MATLAB (R2014a) for preprocessing. A pre-processing step linearly interpolated all spectra to an interval of m/z=0.001 and then subjected the data to median intensity normalization to account for sample to sample variation in overall signal intensity unrelated to molecular patterns. All spectra were peak picked and solvent related peaks were removed from the data set. Log-based transformation was carried out to account for large variability differences between small and large peaks across mass spectra.

Preprocessed data were subjected to multivariate statistical data analysis. Tissue types in each sample and their spatial distribution were pre-determined by examination of each slide-scanned image by an independent histopathologist. PCA was applied to reduce the number of observed variables into a smaller set of principal components. LDA was used for supervised discrimination and classification. Univariate ANOVA analysis was employed to identify discriminatory m/z variables responsible for class separation. Exact m/z values were retrieved from the raw data and these values were used for
compound identification via database search (Metlin).

Figure 12: DESI data analysis workflow. a) Slide scanned image of tissue section with adenocarcinoma, healthy mucosa and stroma annotated by histopathologist. b) Corresponding DESI image with each tissue type now identified based on spectral data. c) Raw DESI spectral data representing each tissue type. d) Supervised Analysis demonstrating differentiation of each tissue type based on DESI mass spectral data.

4.8.5 DESI-REIMS Cross-Platform Workflow

DESI data was processed in centroid mode. Mass range 600-1000 m/z was used for analysis. All annotated pixels from adenocarcinoma and normal adjacent mucosa were selected. Every 3 mass spectra were averaged into 1 mass spectrum to improve signal to noise ratio. Previously pre-processed, centroided corresponding REIMS data was uploaded into Matlab. Mass range 600-1000 was used for further analysis. Mass accuracy was calibrated by linearly shifting all m/z values by +0.05 Da. Each mass spectrum was denoised with a threshold equal to the mean + 7 standard deviations of the mass.
spectrum. For cross platform processing peak annotation was performed using the *in silico* phospholipid database with 8ppm mass deviation window for DESI and 60 ppm for REIMS data. Leave-one-patient-out cross validation was performed and for each cross validation iteration, ANOVA feature selection was carried out with p-value threshold of 0.05.
5. Results

5.1 Instrument Development

Original proof of concept studies of REIMs technology\textsuperscript{213}, based on animal tissue experiments, were carried out in 2009 using home-built Orbitrap ion trap mass spectrometers connected to a modified commercially available electrosurgical device (Radiosurg 2200) with a custom built handpiece. Although data of sufficient quality and intensity was produced during the analysis of liver and lung parenchyma, spleen, skeletal muscle and pancreas tissue, certain tissue types including gastrointestinal tract mucosa produced spectral data of significantly poorer quality and intensity. In order to develop a fully functioning diagnostic medical device, suitable for use within operating theatres in the UK, a different instrument set up was required, with an atmospheric interface capable of providing high quality data from the analysis of all tissue types, which was also suitable for use within the operating theatre. All medical electrical equipment, including high frequency surgical equipment, must be CE marked to indicate compliance with European Medical Device Directives and must meet the minimum safety requirements specified in EN 60601 which outlines the safety, essential performance and electromagnetic compatibility of medical electrical equipment and systems. EN 60601 was first published in 1977 and is regularly updated and restructured, currently including a general standard, and particular standards for individual device classifications. EN 60601-2-2 \textit{Particular requirements for the basic safety and essential performance of high frequency surgical equipment and high frequency surgical accessories} describes the necessary safety requirements for electrosurgical equipment designed for use within the operating theatre.

5.1.2 Modified Electrosurgery Handpiece

The REIMS compatible electrosurgery handpiece used in this study was designed to be as ergonomically and functionally similar to the most widely used commercial product as possible, and to be compatible with the most commonly used electrosurgery generators, to avoid the introduction of further unfamiliar equipment into the operating theatre and to assimilate REIMS analysis into the normal
surgical workflow. The Covidien Valleylab™ electrosurgery system is the most widely used throughout the UK, therefore the custom built handpiece was designed to be compatible with Covidien Valleylab™ generators and to have features as similar to the Valleylab™ button switch monopolar electrosurgical pencil as possible. The hand piece, produced by Medres Ltd (Budapest, Hungary), is ergonomically and functionally similar to this model but has an additional 15mm long piece of stainless steel tubing with a 2mm internal diameter built into the base of the dissection tip. The stainless steel tubing is connected to a central aspiration line, incorporated into the handpiece design to allow aspiration of the smoke plume into a 2 m long, 3mm diameter section of PTFE tubing, used for the evacuation of the aerosol containing gaseous ions from the surgical site to the mass spectrometer. This additional piece of stainless steel tubing does not affect ergonomics or ease of use. As with commercially available products, the tip can be turned 90° within the handpiece to suit the preference of the operating surgeon. To prevent the presence of an additional piece of tubing on the, often overcrowded, operating table the PTFE tubing used to carry the aerosol from the handpiece to the mass spectrometer was connected to the power supply cable for 1.5 m from the end of the handpiece before diverging towards the instrument some distance away from the operating table. Although commercial devices are available for the evacuation of surgical smoke from the field, previous, unpublished, studies have demonstrated that large volumes of dead space in these systems lead to dilution of samples and poor ion transfer. The ion transfer set up describe above was based upon preliminary results described in Balog et al Anal. Chem. 2010, 82, 7343–7350 and Schafer et al Angew Chem Int Ed Engl. 2009;48(44):8240-2. Ergonomics and ion transfer capability of the handpiece were tested by the ex vivo dissection of porcine liver parenchyma. Spectral data of sufficient intensity were produced using the prototype and I did not feel that the additional features impacted upon quality or ease of dissection. This monopolar electrosurgery handpiece, suitable for open surgical dissection was used for the remainder of the experimental work.
5.1.3 Mass Spectrometers and Atmospheric Interface

A collaboration was formed between our research group at Imperial College London and Waters Corporation (Roundthorn Ind Est, Lancashire M23 9LZ), to develop a time-of-flight mass spectrometer which was both REIMS compatible and suitable for intra-operative use.

On 3\textsuperscript{rd} November 2011, myself and 2 colleagues from Imperial College visited Waters Corporation in Manchester to meet with Dr Steve Pringle (Principal Research Scientist) and Dr Mike Morris (Senior Director) and their team. The primary purpose of the visit was to discuss the instrument capabilities required for REIMS analysis and also, the design requirements for a mass spectrometer, which could be used in an operating theatre environment. At the end of this process the team at Waters provided me with a list of information concerning the facilities available in the operating theatre, in addition to the information relating to heat, size, filtration grade, temperature and noise production required to design an instrument suitable for use in a UK operating theatre environment. This information is provided in Appendix 5.

On 24\textsuperscript{th} November 2011 Professor Zoltan Takats visited the Department of Surgery and Cancer, Imperial College London to discuss site preparation, design of proposed lab space, development and delivery of the Synapt and instrument utilisation/experimental design. Myself and the Imperial College team and the team from Waters subsequently undertook multiple visits between the Waters headquarters in Manchester and St Mary’s Hospital, London, to further instrument and laboratory development.
Professor Takats subsequently visited Imperial College regularly, to plan future work and discuss experimental design prior to his appointment to the Department of Computational and Systems Medicine in 2012 where he continued to lead the development of REIMS technology.

Three mass spectrometers with REIMS capability were subsequently developed by this partnership and installed within Imperial College Healthcare NHS Trust. The unique instrument design incorporated a novel atmospheric interface, consisting of a venturi pump to drive gas flow into the instrument via an inlet capillary and a collision surface where large charge particles could be further broken down into small molecules for analysis. Each instrument bore a CE mark indicating conformity the 2006/95/EC Low Voltage Directive and the 2004/108/EC Electromagnetic Compatibility Directive. They were designed to safely allow aspiration of the aerosol to be driven by the central pressurised air supply from the tandem articulated pendant system within the operating theatre, and also to be mobile in order to be moved in and out of clinical areas as required. Replaceable panel type High Efficiency Particulate Air (HEPA) filters were installed to prevent ejection of particulate matter >3 microns into the environment from the instrument ventilation system. The instruments were designed to operate within the designated optimal temperature of operating theatre suites (18-22°C) and to ensure the instrument did not produce heat that would compromise the optimal operating theatre conditions.

**Synapt G2-S**

A Waters Synapt G2-S, customised to meet standards for intra-operative use was the first mass spectrometer to be delivered to Imperial College on 24th April 2012. This early instrument, with additional external PC proved too large and heavy (Width 830 mm, Length 1530 mm, Height 1780 mm, weight 430kg) to be used in the operating theatre environment and was designated a preliminary laboratory based instrument for early phase testing. After several planning meetings with histopathologists, mass spectrometrists and laboratory technicians, I designed a study protocol for ex vivo tissue analysis and distributed this to the lab manager, histopathologists and all team members involved in tissue collection, analysis and data recording. Once the instrument was installed and ready to use, I organized a meeting with the surgeons at Imperial College whose patients would be recruited into the study. Here I informed them of the study start date and presented the study protocol to ensure
that they were aware of when I would be consenting the patients, at what point in the operation I would enter the operating theatre to collect tissue and reassure them that I would use a histologically sound method of tissue sample collection, always in conjunction with a histopathologist, which would not compromise the oncological assessment of their resection specimens. Correspondence between myself and Dr S Pringle (Waters Corporation) in March 2012, outlining the practical aspects of installation of the first REIMS compatible mass spectrometer in a clinical environment is outlined in Appendix 6. Correspondence between myself and the research group and minutes of meeting focusing on the technical aspects of final instrument development, instrument delivery and installation of the Synapt G2-2 are outlined in Appendix 7 and 8.

Figure 14: Customised Waters Synapt g2-s
92 ex vivo samples of colorectal tissue from 34 patients were analysed using the Synapt G2-S. Clinical data is summarized in table 8.

**Table 8: Clinicopathological characteristics of colorectal tissue samples used in testing of the Waters Synapt g2-s**

<table>
<thead>
<tr>
<th></th>
<th>Cancer</th>
<th>NAM</th>
<th>Dysplasia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Samples</strong></td>
<td>35</td>
<td>52</td>
<td>5</td>
</tr>
<tr>
<td><strong>18 male</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>16 female</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anatomy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caecum</td>
<td>5</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Ascending colon</td>
<td>4</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Transverse colon</td>
<td>3</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Descending colon</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Rectosigmoid</td>
<td>12</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>Rectum</td>
<td>9</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td><strong>T stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T0</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Differentiation</strong></td>
<td>18</td>
<td>10</td>
<td>7</td>
</tr>
</tbody>
</table>

The interface produced data with a low signal to noise ratio, which was not of sufficient quality to differentiate between different tissue types. The lipid peak at 885.5 m/z representing the phosphoinositol PI (38:4) – H is reliably seen in REIMS analysis of the vast majority of human tissue samples and its absence from the spectra implies technical failure. This peak was frequently absent from raw mass spectra of human colorectal tissue, when analysed using the Synapt G2-S. Figure 15 demonstrates example raw spectral data from malignant and benign colonic tissue generated using the Synapt G2-S instrument setup, in contrast to a good quality spectrum of colorectal tissue.
The ion optics setup of the Synapt G2-S is shown below (figure 16). This is a complex set up with an ion mobility component (T wave collision cell). Ion mobility separates ions based on mobility (mass, charge, size and shape) prior to pulsed release into the flight tube. This leads to a 2 dimensional separation of charged particles, firstly when travelling through the ion mobility chamber and secondly when ions reach the detector. This can lead to fragmentation of the data and loss of common species including m/z 885.5. These species may still be present but fragmented and represented as fatty acids, however, fatty acid profiles in tissues are more variable and less reliable for tissue identification than the phospholipids which are used in REIMS analysis.
To address the issues with sensitivity and size which precluded use of the Synapt G2-s as an instrument for REIMS analysis, using information gained from the experimental work described above, the Xevo G2-S i-Knife QTof Mass Spectrometry system was developed, as the next generation REIMS compatible mass spectrometer. The instrument is a hybrid, quadrupole, orthogonal-acceleration, time-of-flight mass spectrometer, based upon a commercial Waters Xevo G2-S QTof.

The instrument was modified to make it a single unit with the following improvements:

1. The overall size of the instrument was reduced from 830 mm x 1530 mm x 1780 mm to 690 mm x 1398 mm x 1695 mm and the weight from 430kg to 300kg.

2. All vacuum pumping equipment was incorporated into the main body of the instrument rather than placed on top of the instrument as with the previous design.

3. The host PC was incorporated into the instrument casing and a touch screen monitor was added, from which all functions of the instrument could be controlled. This was achieved by constructing a new chassis, which accommodates the mass spectrometer, vacuum system and...
computer. The chassis comprises a bespoke air-cooling system which also provides sufficient noise and electromagnetic insulation and includes the HEPA filters to eliminate the potential emission of infectious agents into the environment.

4. The instrument was also equipped with wheels, brakes and shock-absorbers for transportation.

5. The venturi pump can be driven by connecting to the medical air supply at a pressure of 4 bar, this can be taken from the main theatre/anaesthetic air supply using a commercial splitter.

To address issues of data fragmentation, the Xevo G2-S i-knife has a simpler ion mobility set up. In keeping with the original design, large externally charged aqueous droplets formed at the point of electrosurgical dissection were allowed to flow through the inlet capillary and hit the collision surface,
where they fragment into smaller charged particles. The charged particles were only permitted to travel through a step wave (which removed contamination and uncharged particles) and a quadropole to guide ions though to the flight tube and detector. Critically, the ion mobility element was removed, to reduce the risk of fragmentation. Due to the special requirements for ion transfer and ion introduction in REIMS, the ion source region of the instrument, originally a Z-Spray source on the Synapt G2-S, was replaced by a new design. A new atmospheric interface was built, using results of gas flow profiles and ion trajectory modeling. REIMS relies on a two-step ionization mechanism comprising the formation of charged clusters at the surgical tissue dissection site, followed by the transfer of these clusters to the mass spectrometer where a surface induced dissociation type separation produces the final detected ions. The novel atmospheric inlet was designed to facilitate these processes. Waters MassLynx software controls the mass spectrometer and acquires, analyzes, manages, and distributes data from the mass spectrometer.

Ion flow in the improved model is as follows:

1. Samples from the tissue being analysed are introduced via a capillary to the collision surface where large aqueous particles are dissociated into smaller, charged particles.

2. The ions pass through the StepWave ion guide, where contamination and uncharged particles are removed, to the quadrupole, where they may be filtered according to their mass-to-charge ratio.

3. The ions pass into the time-of-flight analyzer. A high voltage pulse orthogonally accelerates the ions up the flight tube, where a reflectron reflects them back towards the detector. Ions of different mass-to-charge ratios arrive at the detector at different times. The difference in the arrival times provides the basis for the recorded mass spectrum.

4. The signal from the detector is amplified, digitized, and transmitted to the software

27 samples (19 normal colorectal mucosa, 8 adenocarcinoma) from 14 patients were analysed using the Waters Xevo G2-S QTof i-knife instrumental setup (table 9).
Table 9: Clinicopathological details of colorectal samples used to test the Xevo G2-S QTof i-Knife system

<table>
<thead>
<tr>
<th></th>
<th>Cancer</th>
<th>NAM</th>
<th>Dysplasia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Samples</strong></td>
<td>8</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>7 male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anatomy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caecum</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Ascending colon</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Transverse colon</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Descending colon</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Rectosigmoid</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Rectum</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td><strong>T stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tx</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Differentiation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In comparison to that of the Synapt G2-S, REIMS data was of improved quality, expected phospholipid peaks were present and overall signal intensity was improved, however the data had high levels of noise. Issues with significant mass drift, instrument contamination and background noise led to poor separation of tissue types with an overall accuracy of 67.9% (see figure 18). Mass drift occurs in ToF instruments due to the effect of small environmental changes, such as a small change in room temperature, on the flight tube. This can lead to a single chemical species being recorded at separate m/z values on different days. A batch effect could also be seen, with samples separating based on day of running rather than tissue type (see figure 18). This was due to heavy contamination of the collision surface with debris from previous samples leading to cross contamination of data.
Figure 18: a) Confusion matrix demonstrating poor overall classification accuracy in the differentiation of colorectal adenocarcinoma and normal adjacent mucosa using Xevo G2-S i-knife system. b) PCA plot demonstrates visual clustering of samples (blue = healthy mucosa, pink = adenocarcinoma) analysed on the same day, thought to be due to instrument contamination.
Xevo G2-S i-knife with heated coil interface

To address, the issue of instrument fouling a new atmospheric interface, to be used with the Xevo G2-S QTof i-knife was developed which incorporates a coil of Kanthal, an iron-chromium-aluminium alloy, heated to 800°C into the collision surface to prevent the build up of contaminants and cross contamination of data.

![Image: Original inlet capillary and cold ball collision surface of Xevo G2-S i-knife demonstrating build up of debris on collision surface. Novel heated kanthal coil interface outside and, heated, inside the instrument.](image)

This new interface was designed and tested using porcine colonic mucosa. To establish optimal electrosurgical settings for the analysis of tissues, testing was carried out using electrosurgery cut mode at different power setting (30, 40, 50, 60, 70W). Higher signal intensity and better quality data were produced using the heated coil interface. No discernable difference in signal intensity or data quality were found with different power settings (figure 20) therefore a standardised power setting of 30W was arbitrarily chosen for all future experiments as this power setting is commonly used during surgical procedures.
Figure 20: Raw REIMS data from analysis of porcine colon mucosa sampled using electrosurgery power settings 30-70W. No significant difference was demonstrated in the quality of spectral data produced using different energy levels.
The issue of mass drift was corrected by performing lock mass correction. One or more high intensity peaks within the spectrum, the exact mass of which is known, is annotated and a non-linear shift is performed to adjust the remaining m/z scale against these known peaks. This counteracts the effect of small inaccuracies in m/z value for each peak occurring secondary to small environmental changes. Peaks used in the lock mass correction process were 699.497 PE(34:1) – NH₄, 744.555 PC(36:1) – H and 885.55 (PI 38:4) – H. An automated background subtract function was added to the data analysis process to increase the signal to noise ratio. This function calculates the level of noise based on the distribution of signal intensities within each bin of the spectrum. By comparing these with the average throughout the spectrum, the background noise level is calculated and removed thus improving signal to noise ratio.

The Xevo G2-S i-knife is also a fully integrated system, with heat and noise insulation, which incorporates the mass spectrometer, power supply systems, gas system for Venturi air jet pump and cooling system into a single machine. All functions are controlled by the integrated touch screen computer, which also serves as a results display screen and user interface.

The Xevo G2-S i-knife was tested for functionality within the operating theatres in St Mary’s Hospital prior to starting further analysis of ex vivo human tissue on this instrument. Elements tested in the operating theatre environment included access via available doorways, compatibility with anaesthetic gas supply and adequacy of connectors for gas tubing and gas supply splitters, adequate access to power supply, available space and effect on room temperature. The instrument had been designed to be compatible with the electricity supply in the operating theatres within the hospital as outlined in appendix 5. Modifications to the venturi gas inlet were necessary to allow connection to the anaesthetic gas supply via a ¾ inch diameter gas hose with a Schrader style male quick connect coupler connected to a pendant splitter unit.
5.2 Analysis of Colorectal Cancer Tissue by Rapid Evaporative Ionisation Mass Spectrometry

The Xevo G2-S i-knife system was applied to the ex vivo REIMS analysis of colorectal tissue. Sixty-five tissue samples from 40 patients undergoing colorectal resection at St Mary’s Hospital, London, were collected and analysed. This resulted in the generation of 174 data points. Eighty-one data points representing healthy mucosa from 35 normal tissue samples taken from 28 patients, 79 representing adenocarcinoma from 25 tumour samples from 19 patients (one patient provided samples from 4
synchronous colonic tumours), 14 representing dysplasia from 4 tissue samples from 4 patients (see figure 22).

Six patients (patient ID 15, 44, 62, 63, 68, 72) provided more than one sample of tumour tissue. 7 patients (patient ID 53, 58, 62, 68, 69, 70, 71) provided more than one sample of normal adjacent mucosa. T2 = 8, T3 = 12, T4 = 5 N0 = 11, N1 = 10, N2 = 4. Four metachronous right sided colon cancers were sampled from one patient; of these 3 were T3 and one was T2. For a summary of the clinical data, see table 10.

Figure 22: Summary of healthy, benign and malignant colorectal tissue samples collected and spectral datapoints obtained
Table 10: Clinicopathological characteristics of colorectal tissue samples used to build REIMS database

<table>
<thead>
<tr>
<th>Clinicopathological Characteristic</th>
<th>NAM n=35</th>
<th>Cancer n=25</th>
<th>Dysplasia n = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (Median)</strong></td>
<td>68y (range 47-90)</td>
<td>64 (47-81)</td>
<td>68 (47-90)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td>22 male</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>18 female</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td><strong>Anatomy</strong></td>
<td>Caecum</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Ascending colon</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Transverse colon</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Descending colon</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rectosigmoid</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Rectum</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td><strong>T stage</strong></td>
<td>Tx</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>T1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>T4</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td><strong>N stage</strong></td>
<td>N0</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>N1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>N2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>M stage</strong></td>
<td>M0</td>
<td>32</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>M1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td><strong>Differentiation</strong></td>
<td>Moderate</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td><strong>KRAS</strong></td>
<td>Positive</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td><strong>EMVI</strong></td>
<td>Positive</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td><strong>Type of surgery</strong></td>
<td>Right hemi</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Ext Right hemi</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Sigmoid colectomy</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Anterior Resection</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>APR</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Neoadjuvant therapy</strong></td>
<td>Radiotherapy</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Chemotherapy</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Chemoradiotherapy</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>
5.2.1 REIMS Analysis of the Colorectal Mucosa Lipidome

Figure 23 shows an example of the raw spectra produced by REIMS analysis of normal colonic mucosa and colonic adenocarcinoma. Discriminatory phospholipid peaks within the m/z region 600-900 are highlighted. These include Phosphatidic acids (PA), phosphatidyglycerols (PG), Phosphatidylcholines (PC), Phosphatidylinositol (PI) and Phosphatidylserine (PS), all significant constituents of the mammalian cell membrane lipid bilayer.

![Figure 23: Exemplar raw spectra from REIMS analysis of normal adjacent colonic mucosa and colonic adenocarcinoma demonstrating phospholipid peaks detected in the 600-900 m/z range.](image)

5.2.2 Validating the REIMS Database of Colorectal Adenocarcinoma

Samples taken from surgically resected colorectal adenocarcinoma and normal adjacent mucosa were analyzed ex vivo by REIMS for the construction of a histologically specific mass spectral database for the differentiation of cancer and healthy tissue. Datapoints representing dysplasia were therefore withheld.
from the dataset. Tissue analysis yielded charged molecular species which were subjected to separation based on differing mass-to-charge ratios and subsequent quantitative detection within the mass spectrometer. Each mass spectrum was generated within 0.5s of tissue sampling. Examples of resulting spectra are shown in Figure 23. Ions were predominantly detected in the 600 – 900 m/z range. Unsupervised analysis with PCA demonstrated clustering of diagnostic subgroups (Figure 24a) and supervised analysis with LDA lead to marked separation between the 2 groups (figure 24b). Mahalanobis distance calculations and leave one patient out cross validation were applied. Invasive adenocarcinoma was differentiated from normal adjacent colonic mucosa, based on tissue lipid profiles, with an overall accuracy of 93.6% (figure 24c). Sensitivity 93.6%, specificity 93.8% positive predictive value 93.6% and negative predictive value 93.8%. Area under the curve was 0.97 (figure 24d).
As shown in the confusion matrix (figure 24c), 5 datapoints representing 4 samples of adenocarcinoma were misclassified as healthy mucosa and 5 datapoints representing 5 samples of healthy tissue were misclassified as adenocarcinoma. Misclassified samples are shown in figure 25. The four samples of adenocarcinoma which were misclassified as healthy mucosa were a heterogenous group with no similar clinical features apart from a positive KRAS status. Misclassified samples of adenocarcinoma and normal adjacent mucosa are summarized in table 11.
<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Macroscopic Description (pre-analysis)</th>
<th>Microscopic Description (post-analysis)</th>
<th>Misclassification</th>
<th>Clinical characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 S4 BS01</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Misclassified as NAM</td>
<td>Poorly differentiated mucinous caecal adenocarcinoma arising in villous adenoma. T2 N2. No neoadjuvant therapy. KRAS mutation.</td>
</tr>
<tr>
<td>18 S2 BS01</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Misclassified as NAM</td>
<td>Moderately differentiated rectosigmoid adenocarcinoma arising in adenoma. T3, N1. No neoadjuvant therapy. KRAS mutation</td>
</tr>
<tr>
<td>44 S1 BS01</td>
<td>Normal</td>
<td>Invasive adenocarcinoma</td>
<td>Misclassified as NAM</td>
<td>Moderately differentiated rectal adenocarcinoma arising from villous adenoma. T4 N0. No neoadjuvant therapy. KRAS status unknown</td>
</tr>
<tr>
<td>73 S1 BS01</td>
<td>Polyp</td>
<td>Invasive adenocarcinoma</td>
<td>Misclassified as NAM</td>
<td>Moderately differentiated rectal adenocarcinoma. T2, N0. No neoadjuvant therapy. KRAS mutation</td>
</tr>
<tr>
<td>73 S1 BS02</td>
<td>Polyp</td>
<td>Invasive adenocarcinoma</td>
<td>Misclassified as NAM</td>
<td>Moderately differentiated rectal adenocarcinoma. T2, N0. No neoadjuvant therapy. KRAS mutation</td>
</tr>
<tr>
<td>55 S2 BS01</td>
<td>Normal</td>
<td>Healthy mucosa</td>
<td>Misclassified as adenocarcinoma</td>
<td>Normal rectal mucosa from anterior resection for moderately differentiated rectosigmoid adenocarcinoma. T3, N0. No neoadjuvant therapy.</td>
</tr>
<tr>
<td>65 S3 BS03</td>
<td>Tumour</td>
<td>Healthy mucosa. Scattered tumour cells 2mm from sampling point</td>
<td>Misclassified as adenocarcinoma</td>
<td>Normal mucosa from high anterior resection for poorly differentiated mucinous descending colon adenocarcinoma. T3 N0. No neoadjuvant therapy. KRAS mutation.</td>
</tr>
<tr>
<td>70 S1 BS01</td>
<td>Normal</td>
<td>Healthy mucosa</td>
<td>Misclassified as adenocarcinoma</td>
<td>Normal descending colon mucosa from AP resection for low rectal adenocarcinoma. T3, N0. Neoadjuvant chemo-radiotherapy. KRAS mutation.</td>
</tr>
<tr>
<td>80 S2 BS01</td>
<td>Normal</td>
<td>Healthy mucosa</td>
<td>Misclassified as adenocarcinoma</td>
<td>Normal rectosigmoid from anterior resection for</td>
</tr>
</tbody>
</table>
44 S1 is the most significantly misclassified sample, occupying a space completely within the “normal dataspace” (figure 25). Samples 18 S2 (adenocarcinoma), 55 S2 (NAM) and 80 S2 (NAM) also cluster closely with the opposite sample group. Two samples were obtained from patient number 44 who underwent an anterior resection for what was thought to be a villous adenoma with high-grade dysplasia of the rectum. Postoperatively this was upgraded histologically to invasive adenocarcinoma. Sample 1, although thought to be normal mucosa on sampling, was subsequently found, on post sampling histological validation, to be invasive adenocarcinoma and was therefore labeled as invasive adenocarcinoma for further analysis. Taking into account the initial sample labeling as healthy mucosa, and the degree of misalignment with the adenocarcinoma dataspace, the misclassification may be secondary to sampling error. Samples 55 S2 and 80 S2 were both samples of healthy rectum/rectosigmoid mucosa taken from patients with invasive malignancy who had not been exposed to neoadjuvant therapy. The samples were analysed on separate days. Therefore misclassification was unlikely due to instrument contamination or batch effect. 18 S2 was a sample of moderately differentiated rectal adenocarcinoma from a patient who had not received neoadjuvant therapy. 65 S3 was a mixed sample, on microscopic examination the REIMS sampling point was seen going through normal mucosa with scattered tumour cells 2mm away. Due to tumour cells being more than 2mm from the REIMS sampling point, the sample was labeled as normal mucosa for the purposes of further analysis but was misclassified by the model as adenocarcinoma.
5.2.3 Lipid Profiling of Rectal and Colonic Tissue

Seven out of the ten misclassified colorectal samples (4/5 adenocarcinoma, 3/5 NAM) originated in the rectum or rectosigmoid junction. Rectal and colonic cancers are known to be biologically distinct, and this raised the question of whether the REIMS lipid profile of rectal adenocarcinoma, or even normal rectal mucosa differs from that of the more proximal colon. Tumour samples were separated into those originating in the rectum and rectosigmoid junction (n=12) and those originating in the remainder of the colon (n=13). PCA and LDA, with Mahalanobis distance calculations and leave one sample out cross validation were performed to determine whether the lipid profiles produced during REIMS analysis of rectal adenocarcinoma were sufficiently different from those produced during analysis of more proximal tumours for prospective differentiation within the database.
Adenocarcinoma arising in the rectum and rectosigmoid could be differentiated from more proximal tumours with an overall accuracy of 87.3%. This could not be explained by increased exposure to neoadjuvant therapy in rectal cancer as the database contains no rectal tumour tissue from patients who have undergone neoadjuvant therapy. To determine whether the observed differences could be accounted for by anatomical geographical variation, healthy mucosa samples were separately divided into 2 classes, rectum/rectosigmoid and the remainder of the colon. Although less accurate, it was possible to differentiate these samples with an overall accuracy of 74.1% suggesting some variation in mucosal lipid profiles in different regions of the colon. Gut bacterial colonization and microbiome...
composition change throughout the GI tract. Differences in spectral profiles produced by REIMS analysis of different regions could be secondary to change in bacterial species, as many of the phospholipid species detected are also present in significant numbers in bacterial cell membranes.

5.2.4 Lipid Profiling of Tumour Characteristics

Twenty-five samples of colorectal adenocarcinoma were separated into classes based upon characteristics that influence prognosis and clinical outcome including T stage, N stage, level of differentiation and extramucosal vascular invasion (figure 27).

Figure 27: LDA plots, confusion matrices and ROC curves demonstrating prediction of T stage (T2 vs T3/T4), lymph node involvement and level of differentiation (moderately differentiated vs poorly differentiated) by REIMS

Overall classification accuracy in the identification of tumour tissue by all 3 T stages (T2, T3 and T4) was 69.6%. T2 tumours (n=8) were separated from those T3 and above (n=17) and analysis repeated. T2 and T3/4 tumour samples could be differentiated with an overall accuracy of 85.2%, AUC 0.87. Eight
datapoints representing T3/4 adenocarcinoma were misclassified as T2. Seven of these originated from analysis of the three of the 4 metachronous tumours originating in the right colon of one patient (3 were T3, 1 was T4). The remaining misclassified sample originated from a T3 caecal tumour.

Colonic and rectal adenocarcinoma with lymphnode involvement (n=11) could be differentiated from node negative disease (n=14) with an overall accuracy of 69.9% (AUC 0.77, Sensitivity = 70% and specificity = 68%). Moderately differentiated adenocarcinoma (n=17) could be differentiated from poorly differentiated (n=8) with overall accuracy of 73.4% (AUC=0.80%. Sensitivity for the identification of poorly differentiated adenocarcinoma = 69%, specificity = 77%).

Figure 28: LDA plots, confusion matrices and ROC curves showing more accurate prediction of T stage (T2 vs T3/4), lymph node involvement and level of differentiation by REIMS in rectal adenocarcinoma only
REIMS classification accuracy for the differentiation of tumour tissue by T stage, N stage and level of histological differentiations improve when applied to samples of rectal cancer only (figure 28). Sensitivity and specificity for identification of T3/T4 disease were 92% and 67% respectively. Sensitivity and specificity for the identification of lymphnode positive disease were 70% and 88%. Sensitivity and specificity for the identification of poorly differentiated rectal adenocarcinoma were 67% and 100% when applied to rectal cancer.

Samples of colonic and rectal adenocarcinoma were differentiated based on the presence or absence of a KRAS mutation and EMVI (figure 29). KRAS status was identified from the formal post operative histopathology report following surgical resection and this information was not available for every sample in the dataset. REIMS analysis was accurate in the identification of KRAS mutation within colorectal adenocarcinoma tissue (sensitivity 85%, specificity 82%. AUC 0.95) and those that were positive for EMVI (AUC 0.93, Sensitivity 77%, Specificity 89%. AUC 0.93)
Figure 29: LDA plots, confusion matrices and ROC curves demonstrating prediction of presence of KRAS mutation and extra mural vascular invasion (EMVI) by REIMS

5.2.5 Identifying Dysplasia

14 dysplasia datapoints generated from REIMS analysis of 4 tissue samples from 4 patients were subsequently included in the analysis to determine whether dysplastic tissue could be differentiated from both healthy colorectal mucosa and invasive adenocarcinoma.

A summary of samples of dysplasia samples included in the analysis can be found in Table 12.
Overall accuracy for the identification of malignant, dysplastic and healthy tissue was 83.9% (figure 30). Over 90% of healthy tissue samples were identified correctly, however, 10% of tumour samples were misclassified as dysplasia and 64% of dysplasia samples were misclassified as invasive cancer. The small number of adenocarcinoma samples misclassified as healthy mucosa and vice versa were the same samples that were misclassified in the previous analysis. Most of the reduction in accuracy was a result of misclassification of dysplasia samples as invasive malignancy and to, a lesser degree, vice versa. All datapoints from sample 84 S1 were misclassified. This accounts for 7 of the 9 misclassified dysplasia datapoints. The remaining 2 originated from 44 S2 and 61 S2. There were no clinical features associated with sample 84 S1 which may explain the misclassification. The small sample size in the dysplasia cohort makes further interpretation difficult. Sample 44 S2, although originating from a sample containing invasive malignancy, did not behave differently from other samples of dysplasia.

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Clinicopathological information</th>
</tr>
</thead>
<tbody>
<tr>
<td>37 S2 BS01</td>
<td>Tubulovillous adenoma with low grade dysplasia. No malignancy</td>
</tr>
<tr>
<td>37 S2 BS02</td>
<td>Sample contains dysplastic tissue taken from a lesion which was characterized histologically as moderately differentiated adenocarcinoma arising from a villous adenoma with high grade dysplasia</td>
</tr>
<tr>
<td>44 S2 BS01</td>
<td>Adenomatous polyp with low grade dysplasia. No invasive malignancy.</td>
</tr>
<tr>
<td>44 S2 BS02</td>
<td>Villous adenoma with low grade dysplasia. No invasive malignancy.</td>
</tr>
<tr>
<td>60 S2 BS01</td>
<td></td>
</tr>
<tr>
<td>60 S2 BS02</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS01</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS02</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS03</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS04</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS05</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS06</td>
<td></td>
</tr>
<tr>
<td>84 S1 BS07</td>
<td></td>
</tr>
</tbody>
</table>
Tumour samples were removed from the dataset and healthy mucosa compared directly with samples of dysplasia. In this case it was possible to differentiate these two tissue types with an accuracy of 96.8% as shown in figure 31. Only 1 sample of healthy mucosa was misclassified as dysplasia and 2 sample of dysplasia misclassified as healthy tissue. All healthy mucosa samples were then removed from the dataset and dysplasia and invasive malignancy directly compared. 11 out of the 14 samples of
dysplasia were misclassified into the tumour category whilst only 3 of the malignant samples were misclassified as dysplasia. Due to the large discrepancy in sample numbers in each group this leads to a misleading overall accuracy of 84.9%, however the ROC curve is a more accurate statistical representation of the result with an AUC of 0.75. In order to overcome this discrepancy, tumour and dysplasia were combined into a single class and compared with healthy tissue to determine whether all dysplastic tissue could be differentiated from normal healthy tissue. Differentiation was possible with an overall accuracy of 92% with a sensitivity of 91% and specificity of 93%.

Figure 31: Scores plots, confusion matrices and ROC curves from analysis of NAM and dysplasia, invasive cancer and dysplasia and NAM and all abnormal tissue
5.2.6 Detecting Alterations in the Individual Lipid Species Associated with Colorectal Neoplasia

The distinct mass spectral profile of each tissue types is dependent upon detectable differences in the lipid membrane composition, released during thermal degradation. Kruskal-Wallis analysis of variance (ANOVA) was performed (p<0.05) on a leave one patient out basis and identified significant differences in mean intensities of peaks representing lipid species responsible for the identification of cancerous tissue and normal adjacent mucosa. Exact m/z values were retrieved from the raw data and corresponding candidate lipid species were identified by database search (Metlin). The most statistically significant species are shown in table 13 with m/z value, chemical structure and p value.

Table 13: Peaks most highly significant for the differentiation of cancer and normal adjacent mucosa. The exact mass and glycerophospholipid type of each species are specified below the structure, with the number of carbon atoms and double bonds in parentheses, and an indication of either loss of hydrogen or ammonia indicating how the negative ion is generated.

<table>
<thead>
<tr>
<th>m/z value</th>
<th>Lipid species</th>
<th>Cancer vs. NAM</th>
<th>Dysplasia Vs. NAM</th>
<th>Cancer vs. dysplasia</th>
</tr>
</thead>
<tbody>
<tr>
<td>616.3949</td>
<td>PC(32:0)</td>
<td>1.1e-05 Increase in cancer</td>
<td>4.6e-05 Increased in dysplasia</td>
<td>Not significant p=0.17243</td>
</tr>
<tr>
<td>672.4974</td>
<td>PE(-32:1)</td>
<td>0.002 Increase in cancer</td>
<td>Not significant p=0.69599</td>
<td>Not significant p=0.08</td>
</tr>
<tr>
<td>674.4161</td>
<td>Prostaglandin D2-biotin</td>
<td>4.8e-05 Increase in cancer</td>
<td>Not significant p=0.92472</td>
<td>0.0291 Increase in cancer</td>
</tr>
<tr>
<td>697.4814</td>
<td>PA(36:3)</td>
<td>2.06e-05 Increase in NAM</td>
<td>0.02533 Increased in NAM</td>
<td>Not significant p=0.40809</td>
</tr>
<tr>
<td>718.5392</td>
<td>PE(34:0)</td>
<td>1.7e-06 Increase in cancer</td>
<td>0.00286 Increased in NAM</td>
<td>1.7e-05 Increase in cancer</td>
</tr>
<tr>
<td>751.5283</td>
<td>PA(40:4)</td>
<td>0.016563 Increase in NAM</td>
<td>0.059341 Increased in NAM</td>
<td>Not significant p=0.11928</td>
</tr>
<tr>
<td>769.5025</td>
<td>PG(36:4)</td>
<td>0.00045 Increase in NAM</td>
<td>0.050841 Increased in NAM</td>
<td>Not significant p=0.82987</td>
</tr>
<tr>
<td>844.6</td>
<td>PS(40:1)</td>
<td>5.4e-10 Not significant</td>
<td>Not significant p=</td>
<td></td>
</tr>
</tbody>
</table>
Identified species responsible for tissue discrimination included phosphatidylethanolamines, phosphatidic acids, glycerophospholipids and phosphatidylinositol. The vast majority of these lipid species were detected across all 3 tissue types however, distribution patterns were markedly different. Further ANOVA was performed on the cross validation of invasive cancer and dysplasia and normal adjacent mucosa and dysplasia to determine whether peaks statistically significant in the distinction of adenocarcinoma and healthy mucosa were also significant in the transition from healthy tissue through dysplasia to cancer. The majority of lipid peaks that were significantly increased or decreased in cancer tissue compared to normal adjacent mucosa were also significantly increased or decreased in the same direction in dysplastic tissue. This suggests a spectrum of change from normal adjacent mucosa through dysplasia to malignancy. There are significant differences in the lipid species between normal adjacent mucosa and both dysplasia and malignancy but these same species are not significantly changed in the progression from dysplasia to malignancy, and when they are, continue to be increased in invasive malignancy. Independent peaks associated with dysplasia were also identified on ANOVA performed on the cross validation of invasive cancer and dysplastic samples. Further significant discriminatory lipid peaks include m/z 766.5028 (PS 36:4) (p = 0.00068636), m/z 767.5080 (PI(0-30:0) (p = 0.0022795 ) and 797.6066 PA(22:2/21:0) (p = 0.0015852) with a mean intensity increased in invasive malignancy and
853.4 PI(36:6) (p = 0.0011763), 853.4873 PI(30:6) (p = 0.0011763) and 857.5186 PI(36:4) p = 0.0048805 with a mean intensity increased in dysplasia as compared to invasive malignancy.

Figure 32: ANOVA with example box plots of statistically significant m/z peaks, the associated lipid species and p value for the discrimination of cancer and NAM.

5.3 Analysis of Breast Cancer Tissue by Rapid Evaporative Ionisation Mass Spectrometry

To overcome the significant challenges of working with lipid rich breast tissue, it was necessary to develop a suitable method for REIMS analysis that would yield high quality mass spectral data from analysis of both glandular/fibrofatty breast tissue and neoplastic lesions with a lower lipid content. REIMS analysis of lipid rich tissue is complicated by low yield of phospholipid signal from thermal degradation of tissue with a high fat content which leads to very poor MS signal intensity during analysis of these tissues. Additionally, attempts to improve signal intensity can lead to the production of spectra with high signal intensities in the triglyceride (850-1100 m/z) and fatty acid (150-400 m/z)
regions (Figure 33). Tissue fatty acid and triglyceride profiles to not have the same discriminatory information as phospholipid profiles and do not contribute to differentiation of tissue types, however, these peaks can be of very high intensity and overpower any small phospholipid signal which is produced.

![Spectral data](image)

**Figure 33**: REIMS raw spectral data demonstrating fatty acid peaks at m/z 150-400, phospholipids at 600-850 and triglycerides from 850-1100.

The instrument development strategy was to systematically analyse the influence of ion mode and atmospheric interface configuration to produce discriminatory phospholipid signal from both high lipid normal breast tissue and more fibrous neoplastic breast tissue that could be used in the differentiation of histologically distinct tissue types.

### 5.3.1 Method Development

1. **Negative ion mode REIMS with heated coil interface.**

   Conventional negative ion mode REIMS with a standard atmospheric interface as described in Chapter 3 was used for the analysis of ex vivo breast tissue. Analysis of solid tumor tissue yielded triglyceride signal in the 800-100 m/z range, however analysis of highly adipose normal breast tissue resulted in low intensity data lacking any triglyceride or phospholipid peaks and therefore very low signal intensity (figure 34). Multivariate analysis was not performed to assess the predictive power of these data due to
the very low signal intensity obtained from the analysis of healthy breast tissue.

Figure 34: Spectral data obtained from REIMS analysis of lipid rich healthy breast tissue, demonstrating a few fatty acid peaks only, and of invasive breast cancer demonstrating triglyceride signal.

2. Post ionization REIMS

The technique of post-ionisation was proposed as a potential solution to this problem. Post-ionisation can be used to yield higher signal from phospholipids in fatty tissue by forcing ionization of large lipid droplets by bombarding them with a charged electrospray. An electrospray post ionization REIMS (PI REIMS) atmospheric interface was developed, which utilised an electrospray to promote dissociation and ionisation of the large lipid droplets that are produced on electrosurgical dissection of breast tissue. The solvent (methanol:water + formate) leads to protonation of these large particles, which then dissociate at the collision surface and are transferred into, and analysed by the mass spectrometer in the same way as described for colorectal tissue analysis. The ionisation mechanism is by donation of a proton from formate in the electrospray solvent, however, as protonation gives the particles a positive charge, results are measured in positive ion mode. PI REIMS analysis was carried out on the Waters Xevo G2-S i-knife QTof mass spectrometer. Optimal parameters for post ionization REIMS were based upon results published in Guenther et al/ Electrospray post-ionization mass spectrometry of
electrosurgical aerosols. Optimal instrumental parameters were determined by systematically altering solvent flow rate, electrospray source voltage and nitrogen flow rate until the highest peak intensity was reached.

Instrument setup was as follows:

- Electrospray mounted onto venturi pump
- Solvent: Methanol/water 1:1 + 1% formate
- Solvent flow rate 450microlitre/hr
- Voltage of electrospray source: 4kv
- Nitrogen flow: 7 bar.
- Extrusion distance of the electrospray capillary 5mm.

The decision to use protonisation as the method of ionization followed by analysis in positive ion mode was made based on results by Guenther et al who observed a 20-50 fold signal enhancement using post ionization in positive mode, whilst no improvement was observed in negative mode. In positive ion, mode electrospray post-ionization REIMS analysis resulted in high intensity triglyceride peaks in the m/z range 850-1000 when normal breast tissue was analyzed, but poor signal acquisition during analysis of tumour tissue. This is because the abundant triglycerides in fibrofatty breast tissue readily undergo ionization in positive mode, however most phospholipids do not form positive ions during REIMS analysis, even during electrospray post ionization. Due to the nature of healthy breast tissue being high in diglyceride and triglyceride species, rapid instrument contamination and blocking of the capillary occurred, secondary to build up of lipid material during analysis of samples with such high fat content.
Figure 35: Raw Spectral data acquired from PI REIMS analysis of invasive cancer and healthy breast tissue demonstrating poor signal acquisition and some triglyceride peaks in invasive cancer tissue and high intensity triglyceride peaks in healthy breast tissue.

Detection of tumour tissue in the presence of high lipid normal tissue was therefore poor using this method due to the poor diagnostic power of the overwhelming triglyceride profiles. There was also heavy fouling of the instrument due to analysis of tissue with high fat content and spectral characteristics were heavily influenced by small, immeasurable changes in instrument setup including electrospray voltage, source geometry, flow rate, Venturi inlet pressure and composition of the electrospray solvent. The PI REIMS interface was composed of a free standing electrospray positioned orthogonally to the inlet capillary. The electrospray capillary was led through the Venturi pump and the capillary tip was secured into central position relative to the Venturi housing by a thin fibre. The electrosurgery smoke plume was fed into the electrospray via the handpiece PTFE tubing. The aerosol and electrospray were directed by the nitrogen flow orthogonal to the inlet capillary. Using this set up, even small changes in source geometry significantly influenced spectral characteristics and this was very difficult to standardize.
A more robust configuration was developed to improve reproducibility in the form of a T-piece PI source. This allowed the electrospray to be in direct communication with the inlet capillary, providing more stable source geometry. A novel atmospheric interface featuring a heated jet disruptor surface (heated to 800 °C) was also constructed to reduce contamination of the instrument by the high fat content of normal breast tissue. Using this setup novel setup, the intensity of triglyceride cations in healthy breast tissue increased 2 fold, however lack of discriminatory phospholipid signal and poor signal acquisition from analysis of tumour tissue led to poor tissue discrimination.

3. Isopropanol (IPA)

As previously described, externally charged particles are not formed during REIMS analysis of tissue with a high lipid content. Instead, large lipid droplets are formed which have no external charge and to not interact with the collision surface of the mass spectrometer to produce ionic species which are
suitable for analysis. The aim of PI REIMS was to forcibly ionize these particles by protonation using a formic acid electrospray. This however, only produced signal from readily ionized triglyceride species and not phospholipids. An alternative ionization mechanism was developed. The electrospray element of post-ionisation was removed and isopropanol was added to the nitrogen venturi supply. Use of a solvent leads to dissociation of the large uncharged lipid droplets into smaller charged particles on contact. These particles then dissociate further at the collision surface into measurable small molecules. The heated jet disruptor collision surface element of the PI REIMS set up was maintained to reduce contamination and memory effect. Results were measured in negative ion mode.

Instrument setup was as follows:

- Venturi pump with nitrogen flow of 7 bar
- IPA flow rate 450microlitre/hr
- Venturi outflow in direct communication with inlet capillary of mass spectrometer
- T piece atmospheric Interface and inlet capillary arrangement as per PI REIMS.

This method led to increased detection of phospholipid signal during the analysis of both fibrofatty/glandular breast tissue and solid tumour tissue. Although the overall signal intensity was lower than the intensity of the triglyceride peaks produced by healthy breast issue using PI REIMS, the spectral data produced using this method provided enough discriminatory capacity to differentiate between normal breast tissue and breast cancer tissue. IPA REIMS spectra of normal breast tissue produced spectra with triglyceride peaks with intensity comparable to that of phospholipids, whilst peaks from membrane phospholipids were more dominant in fibroadenomas and breast cancer (figure 37).
### 5.3.2 Prospective Observational Analysis of Diagnostic Sensitivity of REIMS in Breast Cancer

Samples of invasive breast cancer, benign breast lesions and normal healthy breast tissue were collected from patients undergoing breast resection at Charing Cross Hospital, London. Tissue samples were analysed *ex vivo* by REIMS for the construction of a histologically specific mass spectral database for the differentiation of malignant, benign and healthy tissue. One hundred and eleven tissue samples from 45 patients (14 from Imperial College Tissue Bank) were analysed. Forty-three samples of invasive cancer were collected from 19 patients, 52 samples of healthy breast tissue from 26 patients, 16 samples of benign breast lesions from 6 patients and 6 samples with mixed histology from 4 patients (table 14). This led to the generation of 296 spectral datapoints within the MS database (figure 38). Tissue samples were obtained from two male patients undergoing surgery for breast malignancy. Healthy tissue samples were obtained from both, however, tumour tissue samples only obtained from one of these subjects due to very small volume of tumour tissue and risk of disrupting essential histological staging in procuring tissue samples for research. Five patients included in the study underwent neoadjuvant chemotherapy, 1 neoadjuvant radiotherapy and 1 neoadjuvant
chemoradiotherapy. Although tissue was sampled from 4 patients who underwent surgery for excision of DCIS, no DCIS was found in any of the tissue sampled when it was examined microscopically after REIMS analysis. Tissue samples were categorized into one of 5 tissue types based on histological findings on H&E staining and microscopy (Table 15).

Figure 38: Summary of breast tissue samples
Table 14: Clinical characteristics of breast tissue samples

<table>
<thead>
<tr>
<th></th>
<th>Healthy Breast Sample n = 52</th>
<th>Invasive Cancer Sample n=43</th>
<th>Benign Lesion Sample n = 16</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (Median)</strong></td>
<td>69 (range 26-84)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td>M=7</td>
<td>M=3</td>
<td>M=0</td>
</tr>
<tr>
<td></td>
<td>F=45</td>
<td>F=40</td>
<td>F=16</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td>G1 3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>G2 27</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>G3 6</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>N/A 16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>Tumour type</strong></td>
<td>IDC 31</td>
<td>34</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>ILC 5</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>DCIS 16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Fibroadenoma 0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>ER</strong></td>
<td>+ve 30</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-ve 6</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>n/a 16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>PR</strong></td>
<td>+ve 18</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-ve 18</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>n/a 16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>HER2</strong></td>
<td>+ve 8</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>-ve 28</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>n/a 16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td><strong>Type of surgery</strong></td>
<td>Mastectomy 16</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Risk reducing mastectomy 4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Wide local excision/Local excision 32</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td><strong>Neoadjuvant</strong></td>
<td>Radiotherapy 1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 15: Histological classification

<table>
<thead>
<tr>
<th>Histological Classification</th>
<th>Number of tissue samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal fibrofatty/glandular breast tissue</td>
<td>38</td>
</tr>
<tr>
<td>Invasive breast cancer</td>
<td>36</td>
</tr>
<tr>
<td>Fibroadenoma</td>
<td>16</td>
</tr>
<tr>
<td>Fibrous tissue</td>
<td>14</td>
</tr>
<tr>
<td>Mixed sample</td>
<td>6</td>
</tr>
</tbody>
</table>

Samples placed into the fibrous tissue category included benign breast tissue with a fibrous appearance such as scar tissue and post radiotherapy change. Samples placed into the mixed category included those that were predominantly composed of normal breast tissue with a few very scattered tumour cells at the periphery of the sample (n=3), those where the electrosurgery cut was seen to be crossing both healthy and tumour tissue (n=1) and those where the electrosurgery cut was seen to be less than 2mm from tumour tissue (n=2).

5.3.3 Identification of Outliers

Samples taken from surgically resected breast cancer and normal breast tissue were analyzed ex vivo by for the construction of a histologically specific mass spectral database for the differentiation of malignant tissue and normal breast tissue. Tumour and mixed samples, as the sample contains some malignant cells, were classified together as cancer and normal and fibrous breast tissue were classified
together as healthy breast tissue. Benign breast lesions were excluded from the analysis. Supervised and unsupervised multivariate analysis plots are shown in figure 39. Benign and malignant tissue samples were differentiated with an overall classification accuracy of 82.7% (Figure 39). Sensitivity 78%, specificity 85%. AUC = 0.92 (Figure 39).

![PCA and LDA plots of breast cancer tissue (red) and healthy breast tissue (green).](image)

**Overall Classification Accuracy 82.7%**

<table>
<thead>
<tr>
<th>Healthy</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>84.6%</td>
</tr>
<tr>
<td>Cancer</td>
<td>22.1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Healthy</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>162</td>
</tr>
<tr>
<td>Cancer</td>
<td>19</td>
</tr>
</tbody>
</table>

**AUC 0.92323**

Figure 39: PCA and LDA plots of breast cancer tissue (red) and healthy breast tissue (green). Confusion matrix demonstrating classification accuracy of the REIMS model in the identification of breast cancer tissue.

19 of a total of 86 datapoints representing cancer were misclassified as healthy tissue. Nine of these represent mixed samples, where tumour cells were present in the sample but which were not included in the precise sampling point or the sampling point included both normal breast tissue and tumour
cells. These samples were labeled as malignant for preliminary analysis. The most significantly heterogenous misclassified samples included all datapoints from patient 34 S2 (n=8). Microscopically these samples were found to be benign fibrous tissue with very scattered ILC cells at the periphery of the sample. They were labeled as malignant due to the presence of malignant cells in the sample, however they were “misclassified” in the multivariate analysis as normal tissue. The only replicate of one further mixed sample (30 S2D) was also misclassified. Microscopically the entire sampling point was contained within healthy breast tissue with tumour cells 2mm away from the needlepoint.

![Figure 40: Datapoints representing breast cancer clustering with normal healthy breast](image)

The remaining 10 tumour datapoints misclassified as healthy breast tissue were taken from the tissue samples summarized in table 16.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Classification</th>
<th>Clinical Characteristics</th>
<th>Microscopic Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 S4</td>
<td>Invasive Malignancy</td>
<td>ILC G2 ER +ve PR +ve HER 2 -ve</td>
<td>Sample point includes solid fibrous tissue infiltrated by singly scattered and single files lines of neoplastic cells</td>
</tr>
<tr>
<td>TB502035</td>
<td>Invasive Malignancy</td>
<td>IDC G2 ER -ve, PR -ve, HER 2 +ve</td>
<td>Small (3mm) focus IDC within normal breast tissue</td>
</tr>
<tr>
<td>34 S3</td>
<td>Invasive Malignancy</td>
<td>ILC G2 ER+ve PR+ve HER 2 -ve</td>
<td>Sample point going through invasive malignancy.</td>
</tr>
</tbody>
</table>

The 2 remaining mixed samples which were labeled as tumour prior to analysis were classified as malignant by the algorithm. These samples were:

- 47 S1. Microscopically the sample point included dissection of both tumour cells and fibrofatty tissue.
- 0702381A. Microscopically the sample point included fibrofatty breast tissue and a small focus of ILC (<2mm away)

A secondary analysis of the histological slides of these tissue samples was performed by an independent histopathologist to ensure correct assignment of each sample. Each slide was examined in detail and measurements taken of the exact distance between malignant cells and the edge of the electrosurgery cut. Annotated microscopy photographs of 4 of the samples (3 samples were taken from the same block) are shown in figure 41. Although all of the mixed samples were labeled as cancer, the electrosurgery cut only passed directly through tumour cells in 1 sample (47 S1). This sample was classified as tumour by the algorithm. The remaining samples contained tumour cells within varying distance of the sampling point (0.1mm-1.1mm). These distances and relation to REIMS classification are...
demonstrated in figure 41. A secondary cohort analysis was performed directly comparing “mixed” samples with normal healthy breast tissue. Mixed and healthy samples predominantly clustered together, however, 47S1 clustered separately. This is the only sample where the sample point passes directly through dense tumour cells. There is a second clustering along component 2. These 5 datapoints originate from samples of healthy breast tissue taken from a left sided risk-reducing mastectomy following a previous right mastectomy and left WLE and radiotherapy for cancer. Microscopically this sample contained fibrofatty tissue only with no glandular structures seen.

Figure 41: x10 (in addition to standard x10 eyepiece magnification) magnified images of “Mixed” breast tissue samples demonstrating relationship of tumour cells to electrosurgical burn and REIMS classification. A) Small focus of ILC 0.1mm from electrosurgical burn point. The ratio of ILC: fibrous tissue in the sample is very low. B) Small focus of ILC 0.1mm from electrosurgical burn point. C) Dense cancer cells on one side of the electrosurgical burn and adipose tissue (no glands) on the other. Burn travels through both tissue types. D) Focus of IDC 1.1mm from electrosurgical burn point.
Figure 42: LDA plot of healthy breast tissue (red) and mixed samples (purple) demonstrating 47 S1 and 24 S1 as outliers. 47 S1 is image c in figure 41. The electrosurgical burn travels through both dense tumour cells and adipose tissue.

In order to minimize the confounding effects of mixed samples in this preliminary analysis, these were therefore excluded from the dataset and the analysis was repeated.

Figure 43: Confusion matrix and ROC curve demonstrating classification accuracy for the identification of healthy and malignant breast tissue with mixed samples removed from the dataset.

This resulted in the identification of healthy and malignant tissue with an overall classification accuracy of 88.5%, sensitivity 87.1%, specificity 89.1%, PPV 74.4% and NPV 95% (figure 43).

9 tumour datapoints were misclassified as normal tissue. As shown in table 17, 6 of these represented misclassified samples previously identified on the initial analysis. (Samples 40S4 and 34s3).
Table 17: Misclassified breast cancer samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Classification</th>
<th>Clinical Characteristics</th>
<th>Microscopic Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 S4 (Previously misclassified)</td>
<td>Invasive Malignancy</td>
<td>ILC G2 ER +ve PR +ve HER 2 -ve</td>
<td>Sample point includes solid fibrous tissue infiltrated by singly scattered and single files lines of neoplastic cells</td>
</tr>
<tr>
<td>34 S3 (Previously misclassified)</td>
<td>Invasive Malignancy</td>
<td>ILC G2 ER +ve PR +ve HER 2 -ve</td>
<td>Sample point going through invasive malignancy.</td>
</tr>
<tr>
<td>47 S1 (Newly misclassified)</td>
<td>Invasive malignancy</td>
<td>IDC G2 ER +ve PR +ve HER2 -ve</td>
<td>Specimen consists almost entirely of dense nodules of malignant cells separated by bands of fibrous connective tissue.</td>
</tr>
<tr>
<td>TBCX13000186FT2 (Newly misclassified)</td>
<td>Invasive malignancy</td>
<td>G3 IDC ER -ve PR -ve HER2 +ve</td>
<td>Specimen also shows florid unusual epithelial hyperplasia and a benign calcified lesion</td>
</tr>
</tbody>
</table>

Kruskal-Wallis analysis of variance (ANOVA) was performed (p<0.05) on a leave one patient out basis to identify significant differences in mean intensities of peaks representing lipid species responsible for the identification of cancerous tissue and normal healthy breast tissue. The most statistically significant species are shown in table 18 with m/z value, chemical structure and p value.

Table 18: Most statistically significant m/z peaks responsible for the differentiation of normal and malignant breast tissue, the lipid species, chemical structure and p value

<table>
<thead>
<tr>
<th>m/z value</th>
<th>Lipid species</th>
<th>Cancer vs. Healthy Breast Tissue. P value</th>
<th>Cancer vs. fibroadenoma P value</th>
<th>Healthy vs. fibroadenoma P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>640.475</td>
<td>ClCer(30:2)</td>
<td>1.8e-6 Increased in healthy tissue</td>
<td>Not significant p=0.61225</td>
<td>0.000932 Increased in healthy tissue</td>
</tr>
<tr>
<td>Mass</td>
<td>Compound</td>
<td>p-Value</td>
<td>p-Value</td>
<td>p-Value</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>642.495</td>
<td>GlcCer(30:1)</td>
<td>7.3e-9</td>
<td>Increased in healthy tissue</td>
<td>Not significant</td>
</tr>
<tr>
<td>698.515</td>
<td>PE(34:2p)</td>
<td>7.7e-7</td>
<td>Increased in healthy tissue</td>
<td>Not significant</td>
</tr>
<tr>
<td>703.515</td>
<td>PA(36:0)</td>
<td>Increased in cancer</td>
<td>Not significant</td>
<td>p=0.93514</td>
</tr>
<tr>
<td>717.55</td>
<td>PE-Cer(d37:1)</td>
<td>1.89e-6</td>
<td>Increased in cancer p=0.00021</td>
<td>Not significant</td>
</tr>
<tr>
<td>729.57</td>
<td>PA(39:0)</td>
<td>2.8e-6</td>
<td>Increased in cancer</td>
<td>p=0.035058</td>
</tr>
<tr>
<td>730.53</td>
<td>PC(32:1)</td>
<td>0.007</td>
<td>Increased in cancer</td>
<td>Not significant</td>
</tr>
<tr>
<td>749.535</td>
<td>PA(40:5)</td>
<td>0.0006</td>
<td>Increased in healthy tissue</td>
<td>Not significant</td>
</tr>
<tr>
<td>750.5443</td>
<td>PE(38:4)</td>
<td>1.67e-8</td>
<td>Increased in healthy tissue</td>
<td>Not significant</td>
</tr>
<tr>
<td>758.57</td>
<td>PC(34:1)</td>
<td>7.3e-9</td>
<td>Increased in cancer</td>
<td>Not significant</td>
</tr>
<tr>
<td>765.495</td>
<td>PI(30:1)</td>
<td>2.82e-5</td>
<td>Increased in cancer p=0.000613 Increased in cancer</td>
<td>Not significant</td>
</tr>
<tr>
<td>770.5705</td>
<td>PC(35:2)</td>
<td>0.0009</td>
<td>Increased in healthy tissue</td>
<td>Not significant</td>
</tr>
<tr>
<td>797.55</td>
<td>PG(38:4)</td>
<td>5.28e-5</td>
<td>Increased in cancer</td>
<td>Not significant</td>
</tr>
<tr>
<td>809.5549</td>
<td>PI(33:0)</td>
<td>6.91e-9</td>
<td>Increased in cancer</td>
<td>p=0.012627 Increased in cancer</td>
</tr>
<tr>
<td>863.5655</td>
<td>PI(36:1)</td>
<td>7.216e-8</td>
<td>Increased in cancer</td>
<td>p=0.021466 Increased in cancer</td>
</tr>
<tr>
<td>864.5760</td>
<td>PS(42:5)</td>
<td>0.005</td>
<td>Increased in cancer</td>
<td>Not significant</td>
</tr>
<tr>
<td>885.5499</td>
<td>PI(38:4)</td>
<td>8.5853e-8</td>
<td>Increased in healthy tissue</td>
<td>Not significant</td>
</tr>
<tr>
<td>886.5604</td>
<td>PS(44:8)</td>
<td>1.6918e-8</td>
<td>Increased in</td>
<td>Not significant</td>
</tr>
</tbody>
</table>
Figure 44: ANOVA plot demonstrating peaks representing the monounsaturated PCs, PA, PI and PE which are statistically significant in the distinction of breast cancer and healthy breast tissue.

5.3.4 Identification of Benign Breast Lesions

Multivariate analysis of REIMS spectral data from breast tissue was repeated with 6 samples of benign breast lesion included in the analysis to investigate whether REIMS technology is capable of differentiating between benign, malignant and normal healthy breast tissue. All benign lesions sampled were fibroadenoma, the most common benign breast tumour, neoplastic in nature, composed of epithelial and stromal components and oestrogen responsive ref. Supervised and unsupervised
multivariate analysis plots and a confusion matrix demonstrating the predictive power of the model are shown in figure 45.

![PCA and LDA plots](image)

**Figure 45**: PCA and LDA plots of healthy breast tissue (green), invasive malignancy (red) and fibroadenoma (purple). Confusion matrix showing classification accuracy for the ex vivo identification of healthy breast tissue, invasive cancer and fibroadenoma.

The statistical accuracy of REIMS for the diagnosis of benign tumours in the general model was low however invasive malignancy, benign lesions and normal healthy breast tissue were identified with an overall classification accuracy of 81.3%. A high proportion of healthy tissue samples (89.4%) were identified as such and a high proportion of invasive cancer samples 72.9% were also identified correctly, however, almost half of benign lesions were misclassified as normal tissue (47.4%) although the
number of benign tissue samples was small and therefore misclassification of only a few samples had a high impact on overall accuracy.

To further investigate the similarities and differences between benign lesions and healthy tissue/invasive malignancy, tumour and benign tissue samples from 25 patients were compared. Unsupervised and supervised multivariate analysis plots and confusion matrix are shown in figure 46. In this case the model had a high sensitivity for the detection of malignancy (84%) but only 63.2% of benign lesions were accurately identified. Similarly, direct comparison was made between benign lesions and healthy tissue from 30 patients. Benign lesions were identified with a sensitivity of 79% and specificity of 97%. The confusion matrix and ROC curve can be found in figure 46.

![Confusion matrix and ROC curve demonstrating accuracy of the model for the classification of invasive cancer and fibroadenoma (top) and healthy breast tissue and fibroadenoma.](image)
5.3.5 Tumour Characteristics.

Tumour samples were divided into classes based upon characteristics that influence prognosis and clinical outcome including type of tumour, tumour grade, hormone and HER2 receptor status. Raw spectral data was pre-processed and subjected to unsupervised and supervised analysis using PCA and LDA as previously described. The resulting LDA space was used for classification of spectra from unknown tissue samples using Mahalanobis distance calculations.

Two histological breast cancer subtypes were present in the dataset, invasive ductal carcinoma and invasive lobular carcinoma. ILC is less common than IDC, comprising around 10% of histological subtypes, there are many pathological and clinical features that distinguish these 2 tumour types. ILC is characterised histologically by small round or spindle-shaped cells that grow in a characteristic single-file pattern to infiltrate the adjacent stroma. It is less likely to present as an easily palpated discrete mass as such and often presents at a more advanced stage than IDC. REIMS analysis of 9 samples (15 datapoints) of ILC and 34 (55 datapoints) of IDC correctly identified 100% of IDC samples and 66.7% of ILC samples leading to an overall classification accuracy of 92.9%. The LDA plot and ROC curve are shown in figure 47.

Figure 47: LDA scores plot and ROC curve of REIMS analysis of breast IDC and ILC.
A model was built to determine the accuracy of REIMS analysis in the identification of HER2 overexpression in breast cancer tissue samples. REIMS analysis correctly identified 88.9% (48/54) of HER2 negative samples and 75 % (12/16) HER2 positive samples. The overall classification accuracy was 85.7% (LDA plot, confusion matrix and ROC curve shown in figure 48). Kruskal-Wallis ANOVA with a P value of 0.05 was performed to determine which lipid peaks were statistically significant in the differentiation of the 2 tissue types. Increased levels of certain lipid species, including PS and PI were the most statistically significant in the identification of HER2 positive tumours, however, increased levels of amino acids (Trp Arg Phe Thr/Trp Tyr Lys Gln) were found to be the most discriminatory in the identification of HER2 negative tumours.

![Image of LDA plot, confusion matrix and ROC curve demonstrating accuracy in the differentiation of HER2 positive and negative tumour samples](image)

**Figure 48**: LDA plot, confusion matrix and ROC curve demonstrating accuracy in the differentiation of HER2 positive and negative tumour samples
Tumour samples were divided into ER positive and negative and a model was built to determine whether REIMS analysis could accurately differentiate between tissue types. ER status was reported using the Allred score which combines an estimated proportion of cells that stain positive for ER on a scale of 0-5 with an average intensity score of 0-3. Each sample is then graded on a scale of 0-8. 0-1 indicates a completely negative result (less than 1% of cells are ER positive) and 2-8 is used as a means of semi quantifying the level of immunoreactivity. For the purpose of this analysis, all samples with an ER score of 2 or above were labeled as positive. REIMS performed less well, as a diagnostic test, in the differentiation of ER positive and negative tumour samples than in distinguishing tumour type or HER2 receptor status. Overall classification accuracy was low at 72.9% with 75% of ER positive tumours correctly identified but only 58.3% of ER negative tumours identified as such. The number of oestrogen receptor negative samples in this preliminary dataset was small, as negative tumours represent a small percentage of cases of breast cancer, however this discrepancy in numbers between the two groups should not affect classification accuracy in this type of multivariate analysis and the LDA plot (figure 49) does not demonstrate significant separation of the two tissue types. The ER negative samples misclassified as ER positive scored 0/8 on the Allred scoring system. Samples from 2 patients with ER negative breast cancer were classified correctly and samples from 3 patients were classified incorrectly. There were no distinctive clinical characteristics within these 2 groups which could have contributed to clustering secondary to other features. All samples were ER and PR receptor negative. Correctly classified samples are illustrated on the scores plot. One sample was HER2 receptor positive and one negative, one was grade 2 and one grade 3 and only one patient (31) had neoadjuvant chemotherapy. Within the misclassified sample set there were 2 grade 2 and one grade 3 tumour, and 2 HER2 positive and one triple negative tumour.
Figure 49: LDA plot with correctly classified samples illustrated, ROC curve and confusion matrix demonstrating overall accuracy in the discrimination of ER positive and negative breast cancer

Breast cancer grade is determined using the Nottingham (Elston-Ellis) modification of the Scarff-Bloom-Richardson grading system, also known as the Nottingham Grading System. Tumours are graded based on histological features identified during microscopic examination. These include the degree of tubule or gland formation, nuclear pleomorphism and mitotic count. Each category receives a score between 1 and 3 (1 having appearance closest to that of normal cells and 3 being the least well differentiated). Each tumour is then scored by the histopathologist out of a total of 9 and graded G1 (score 3-5), G2 (score 6-7), G3 (score 8-9). Histological grade represents the morphological assessment of tumor biological characteristics and is therefore a subjective assessment made by experienced clinicians. Breast cancer samples within the dataset were separated into 3 groups based on histological grade (Nottingham Grading System) to determine whether there are distinct, measurable biological
characteristics associated with the intrinsic histological features which determine tumour grade and whether these can be detected by REIMS analysis.

![LDA plot and confusion matrix produced by REIMS analysis of grade 1(green), grade 2 (orange) and grade 3(red) breast cancer samples.](image)

**Overall Classification Accuracy 72.9%**

Figure 50: LDA plot and confusion matrix produced by REIMS analysis of grade 1(green), grade 2 (orange) and grade 3(red) breast cancer samples.

REIMS analysis accurately differentiates grade 1 from grade 3 disease and vice versa. No grade 1 datapoints were misclassified as grade 3 or grade 3 as grade 1. 40% of grade 1 cancers were misclassified as grade 2 and 33% of grade 3 were also misclassified as grade 2 (figure 50). The LDA plot
demonstrates a clear transition from grade 1 through grade 2 to grade 3 demonstrating that there are some objective biochemical changes associated with increasing tumour grade which can be detected in the chemical profiles produced during REIMS analysis.

5.4 DESI/REIMS Cross Validation

As described, DESI MSI enables the spatially resolved simultaneous determination of hundreds to thousands of mall metabolites and complex lipids within a tissue sample by automated sampling of the tissue surface. The ionization process is carried out by directing a pneumatically assisted electrospray of solvent onto the tissue surface. The impact of the multiply charged solvent droplets on the surface results in the formation of secondary charged droplets which contain molecular species from the tissue surface. These species are converted into gaseous ions when the solvent evaporates, which can then by analysed by the mass spectrometer\textsuperscript{218}. Large scale, pre-built, histologically assigned spectral databases are required for REIMS identification of unknown tissue types in near real time. REIMS is an inefficient method for this in terms of sample utilization, relatively poor spatial resolution and potential loss of essential histological information due to its destructive nature. A significant advantage of DESI MSI is that it provides large volumes of data with reliable spatial localization of chemical species within small tissue samples at a high resolution, and is largely nondestructive, leaving intact tissue behind for full histological examination. A potential solution is the utilization of DESI datasets, acquired at a high spatial resolution, to populate the large, histologically assigned spectral databases required for in-vivo REIMS based tissue identification.

DESI MSI detects similar chemical species to REIMS, however, the differences between ionization methods and instrumental platforms introduce platform specific variation into the datasets making direct comparison impossible and significantly hindering the construction of a REIMS tissue identification system using data collected by DESI MSI. The DESI ionization mechanism involves analyte extraction from the tissue surface which leads to the formation of multiply charged droplets which undergo an electrospray-like ionization; in REIMS the tissue components themselves are converted into an aerosol by Joule-heating. In general DESI tends to produce molecular ions and adducts, while REIMS
produces molecular ions of intact molecules and thermal degradation products. However significant overlap within the chemical species detected by each ionization technique is the basis for successful cross-platform normalization. Although mass spectrometric tissue profiling techniques produce different raw mass spectral patterns there is substantial overlap regarding the chemical species detected by the different techniques, (the overlap in chemical species detected using REIMS and DESI exceeds 60%) however, due to molecular ion pattern differences between these platforms the spectra initially need to be preprocessed to obtain chemical profiles that may be comparable.

5.4.1 DESI Analysis of Colorectal Tissue

Prior to attempting DESI REIMS cross platform data analysis, prospective DESI databases of colorectal tissue were built to determine whether DESI spectral data measured using the instrument set up described was sufficient for the classification of histologically distinct tissue types.

28 tissue samples (19 NAM, 8 adenocarcinoma and 1 dysplasia) were analysed by DESI MSI. 21 samples were taken from patients undergoing colorectal resection within Imperial College Healthcare NHS Trust and 7 from the Imperial College Tissue Bank. 21 corresponding tissue samples were analysed by REIMS. (5 adenocarcinoma, 1 dysplasia, 16 normal adjacent mucosa). Clinicopathological features are summarized in table 19. One sample of adenocarcinoma also contained dysplastic tissue and this was annotated and labeled as such.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Sample number</th>
<th>Fresh/bank</th>
<th>Pathology</th>
<th>Location (Surgery)</th>
<th>T stage</th>
<th>N stage</th>
<th>EMVI</th>
<th>KRAS</th>
<th>Diff</th>
<th>Neoadjuvant therapy</th>
<th>REIMS sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>IKCR15</td>
<td>1</td>
<td>Fresh</td>
<td>NAM</td>
<td>Caecum (R hemi)</td>
<td>2</td>
<td>2</td>
<td>Pos</td>
<td>Mut</td>
<td>Poor</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR46</td>
<td>3</td>
<td>Fresh</td>
<td>NAM</td>
<td>Sigmoid (AR)</td>
<td>4</td>
<td>2</td>
<td>Pos</td>
<td>Un</td>
<td>Mod</td>
<td>Yes (chemo)</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR47</td>
<td>2</td>
<td>Fresh</td>
<td>NAM</td>
<td>Rectum (LAR)</td>
<td>2</td>
<td>0</td>
<td>Neg</td>
<td>Un</td>
<td>Poor</td>
<td>Yes (chemo/rad)</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR54</td>
<td>1</td>
<td>Fresh</td>
<td>NAM</td>
<td>Caecum (Ileocaecal resection)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Case No.</td>
<td>Tissue Type</td>
<td>Onset</td>
<td>NAM</td>
<td>Site</td>
<td>Stage</td>
<td>Res.</td>
<td>Adj.</td>
<td>Chemotherapy</td>
<td>Radiotherapy</td>
<td>Final Outcome</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>-------</td>
<td>-----</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>-----</td>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>IKCR54</td>
<td>3</td>
<td>Fresh</td>
<td>NAM</td>
<td>Caecum (ileocaecal resection)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>IKCR57</td>
<td>1</td>
<td>Fresh</td>
<td>NAM</td>
<td>Sigmoid (sigmoid colectomy)</td>
<td>4</td>
<td>1</td>
<td>Neg</td>
<td>Mut</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR59</td>
<td>2</td>
<td>Fresh</td>
<td>NAM</td>
<td>Rectum (AP resection)</td>
<td>2</td>
<td>1</td>
<td>Neg</td>
<td>Neg</td>
<td>Mod</td>
<td>Yes (short course radio)</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR62</td>
<td>4</td>
<td>Fresh</td>
<td>NAM</td>
<td>RSJ (high AR)</td>
<td>3</td>
<td>0</td>
<td>Neg</td>
<td>Mut</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR66</td>
<td>4</td>
<td>Fresh</td>
<td>NAM</td>
<td>Rectum (AR)</td>
<td>3</td>
<td>0</td>
<td>Pos</td>
<td>Mut</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR60</td>
<td>2</td>
<td>Fresh</td>
<td>Dysplasia</td>
<td>Sigmoid (AR)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>IKCR53</td>
<td>4</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>RSJ (AR)</td>
<td>3</td>
<td>0</td>
<td>Pos</td>
<td>Neg</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR61</td>
<td>1</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>Rectum (AR)</td>
<td>3</td>
<td>0</td>
<td>Neg</td>
<td>Mut</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR62</td>
<td>2</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>RSJ (high AR)</td>
<td>3</td>
<td>0</td>
<td>Neg</td>
<td>Mut</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR66</td>
<td>1</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>Rectum (AR)</td>
<td>3</td>
<td>0</td>
<td>Pos</td>
<td>Neg</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IKCR16</td>
<td>2</td>
<td>Fresh</td>
<td>NAM</td>
<td>Caecum (R hemi)</td>
<td>3</td>
<td>2</td>
<td>Pos</td>
<td>Neg</td>
<td>Poor</td>
<td>Yes (chemo)</td>
<td>No</td>
</tr>
<tr>
<td>IKCR21</td>
<td>1</td>
<td>Fresh</td>
<td>NAM</td>
<td>Transverse colon (ext R hemi)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IKCR21</td>
<td>2</td>
<td>Fresh</td>
<td>NAM</td>
<td>Transverse colon (ext R hemi)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IKCR28</td>
<td>3</td>
<td>Fresh</td>
<td>NAM</td>
<td>Caecum (R Hemi)</td>
<td>1</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>IKCR11</td>
<td>3</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>RSJ (high AR)</td>
<td>4</td>
<td>1</td>
<td>Un</td>
<td>Mut</td>
<td>Poor</td>
<td>Yes (chemo)</td>
<td>No</td>
</tr>
<tr>
<td>IKCR12</td>
<td>3</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>Rectal (AP resection)</td>
<td>3</td>
<td>0</td>
<td>Pos</td>
<td>Neg</td>
<td>Mod</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>IKCR51</td>
<td>1</td>
<td>Fresh</td>
<td>Adenocarcinoma</td>
<td>Sigmoid (high AR)</td>
<td>3</td>
<td>0</td>
<td>Neg</td>
<td>Neg</td>
<td>Mod</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>302091</td>
<td>A Bank</td>
<td>NAM</td>
<td>Caecum (right hemicolectomy)</td>
<td>4</td>
<td>2</td>
<td>Un</td>
<td>Un</td>
<td>Mod</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>602011</td>
<td>B Bank</td>
<td>NAM</td>
<td>Descending Colon (Left hemi)</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>602055</td>
<td>B Bank</td>
<td>NAM</td>
<td>Descending Colon (Left hemi)</td>
<td>3</td>
<td>1</td>
<td>Pos</td>
<td>Un</td>
<td>Poor</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>602113</td>
<td>A Bank</td>
<td>NAM (dysplastic)</td>
<td>Right</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Mass spectral images and histopathology images were co-aligned. Morphological features of individual tissue types within each sample were annotated on the slide scanned image by a histopathologist and feature co-registration was carried out to define corresponding regions in the MS image.

Regions corresponding to mucosa, submucosa and smooth muscle were selected from 19 samples of normal healthy colorectal tissue. Lipid profiles representing each pixel in the selected region were generated and MS image reconstruction was carried out based on this profile data. Spectral profiles from each morphologically distinct tissue layer (m/z 600-900) were subject to multivariate analysis by PCA, LDA and Mahalanobis distance calculations as previously described.
Figure 51: DESI MSI analysis of individual tissue layers of normal colonic wall. Unsupervised PCA analysis and confusion matrix.
The same process was repeated choosing regions corresponding to adenocarcinoma and normal healthy mucosa (from both tumour and healthy tissue samples). Differentiation of adenocarcinoma and healthy mucosa was possible with an overall classification accuracy of 93.8% and identification of adenocarcinoma with a sensitivity and specificity of 91% and 96% respectively. The majority of misclassified samples originated from sample 66 S1 (T3 N0 moderately differentiated rectal adenocarcinoma. EMVI positive KRAS negative. No neoadjuvant therapy).

Figure 52: DESI colorectal adenocarcinoma and healthy mucosa

Kruskal-Wallis ANOVA (p value 0.05) was performed to identify spectral peaks significant in the distinction of adenocarcinoma and healthy mucosa by DESI analysis. Actual m/z values were retrieved from the raw data and candidate lipid species identified by database search (Metlin).

M/z values, lipid species and p value are shown in table 20. Seven lipid species significantly altered between the two tissue types in DESI analysis were also significant in REIMS analysis.
Table 20: Statistically significant peaks in the DESI analysis of colorectal cancer with lipid specie and p value

<table>
<thead>
<tr>
<th>m/z value</th>
<th>Lipid species</th>
<th>Cancer vs. Normal Colonic Mucosa. P value</th>
<th>Also significant in REIMS analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>634.4089</td>
<td>(PS(P-26:0))</td>
<td>p = 0.017 Increased in healthy mucosa</td>
<td>No</td>
</tr>
<tr>
<td>688.4923</td>
<td>PE(32:1)</td>
<td>p = 0.013 Increased in adenocarcinoma</td>
<td>No</td>
</tr>
<tr>
<td>672.4974</td>
<td>PE(-32:1)</td>
<td>p = 0.021 Increased in adenocarcinoma</td>
<td>Yes Increased in adenocarcinoma</td>
</tr>
<tr>
<td>699.4970</td>
<td>PA(36:2)</td>
<td>p = 0.014 Increased in adenocarcinoma</td>
<td>No</td>
</tr>
<tr>
<td>714.5079</td>
<td>PE(34:2)</td>
<td>p = 0.042641 Increased in healthy mucosa</td>
<td>No</td>
</tr>
<tr>
<td>744.5549</td>
<td>PC(33:1)</td>
<td>p = 0.009468 Increased in adenocarcinoma</td>
<td>No</td>
</tr>
<tr>
<td>769.5025</td>
<td>PG(36:4)</td>
<td>p = 0.01 Increased in healthy mucosa</td>
<td>Yes Increased in healthy mucosa</td>
</tr>
<tr>
<td>797.5338</td>
<td>PG(38:4)</td>
<td>p = 0.03 Increased in adenocarcinoma</td>
<td>No</td>
</tr>
<tr>
<td>844.6073</td>
<td>PS(40:1)</td>
<td>p = 0.01 Increased in healthy mucosa</td>
<td>Yes Increased in healthy mucosa</td>
</tr>
<tr>
<td>856.5498</td>
<td>PS(P-43:8)</td>
<td>p = 0.02 Increased in adenocarcinoma</td>
<td>No</td>
</tr>
<tr>
<td>885.5499</td>
<td>PI(38:4)</td>
<td>p = 0.005 Increased in adenocarcinoma</td>
<td>Yes Increased in Adenocarcinoma</td>
</tr>
</tbody>
</table>
### Table

<table>
<thead>
<tr>
<th>Mass</th>
<th>PI</th>
<th>p Value</th>
<th>Increased in adenocarcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td>887.5655</td>
<td>PI(38:3)</td>
<td>0.005</td>
<td>Yes</td>
</tr>
<tr>
<td>889.5777</td>
<td>PI(38:2)</td>
<td>0.02</td>
<td>Yes</td>
</tr>
<tr>
<td>911.5655</td>
<td>PI(40:5)</td>
<td>0.05</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Other authors have found a demonstrable difference in the spectral profiles of cancer adjacent mucosa and distant healthy mucosa, a difference attributed to the peritumoural field effects or ‘metaplasia’\(^{151}\). Only 2 cancer samples within this sampleset contained any morphologically normal mucosa. Both samples originated from the rectosigmoid junction. Two corresponding samples of distant healthy mucosa originating from the rectosigmoid junction were selected. A model was built using 50 representative spectra from each of the four samples. It was not possible to differentiate between tumour adjacent and distant mucosa with this very small sample set. PCA analysis resulted in 4 distinct dataclouds corresponding to individual patients and therefore samples classified by patient.

Two samples of dysplasia were present in the dataset, one a sample of adenoma with low-grade dysplasia and one, a sample of peritumoural dysplastic tissue identified on histological annotation of samples. No datapoints from either of these samples were identified as dysplasia by the model. Unlike REIMS analysis, the majority (95%) of dysplasia data were classified as healthy mucosa.
Figure 53: DESI analysis of colorectal adenocarcinoma, dysplasia and normal healthy mucosa.

Overall Classification Accuracy 87.6%
5.4.2 Testing of the DESI REIMS Data Cross Validation Algorithm

No data fusion algorithms exist for the reduction of platform specific variation and combining mass spectral data obtained by different ionization methods and mass spectrometers. The proposed data fusion method requires data pre-processing in order to obtain a summary level of molecular ion species per constituent, followed by application of cross-platform normalization procedures to bring analyte intensities from both platforms to a common scale. The workflow is divided into training and application phases. The training phase requires the acquisition of mass spectra of the same biological sample set on two or more distinct analytical platforms. The outcome of the training phase is a vector of analyte-specific normalization factors that account for intensity differences between analytical platforms. This normalization vector is subsequently used in the application phase to transform analyte intensities from the test platform to a common scale of the reference platform. This method reduces MS platform specific variation allowing clear clustering of metabolite profiles according to histological tissue type, irrespective of MS platform used.

Overall Classification Accuracy 82.25%

Figure 54: Confusion matrix DESI REIMS cross validation
Spectral data representing colorectal adenocarcinoma and normal adjacent mucosa, as measured by REIMS, could be differentiated based on a training set of data produced by DESI analysis of colorectal tissue with an overall accuracy of 82% (figure 54). Sensitivity and specificity for the detection of colorectal adenocarcinoma were 75% and 94% respectively. Lipid species significant in both the DESI and REIMS discrimination of colorectal adenocarcinoma and normal adjacent mucosa include PE(-32:1), PE(34:0), PI(38:3), PI(38:2), PI(40:5) increased in adenocarcinoma and PG(36:4), PS(40:1) increased in normal adjacent mucosa.

5.5.0 iEndoscope

A conventional endoscopic electrosurgical device was adapted to enable the efficient aspiration of the ionic aerosol required for REIMS analysis by addition of several puncture holes in the tubing surrounding the “active” end of the snare to allow improved efficiency of aspiration. To enable mass spectrometric analysis of the electrosurgical aerosol, a BD Connecta™ three-way stopcock (Becton Dickinson UK Ltd) approximately 20 centimetres distal to the snare handle. The snare was then coupled to the mass spectrometer via 0.5 metres of PTFE tubing placed into the third arm of the stopcock. The modified endoscopic snare system was passed down a standard commercially available colonoscope and tissue samples were analysis by touching them with the tip of the snare. MS analysis was performed on the Waters Xevo G2-S QTof i-Knife.
The system was tested for signal intensity and quality using samples of porcine colonic mucosa at a range of power settings (figure 55). High quality spectral data a good signal-to-noise ratio were produced by analysis of porcine colonic mucosa using this method, however overall signal intensities were approximately one order of magnitude lower. Some loss of signal intensity would be expected with the increased distance of the dissection point from the mass spectrometer associated with use of an endoscopic snare.

Figure 55: Raw spectral data produced during analysis of porcine colon with the REIMS endoscopic snare system at a range of power settings.
5.5.1 Colonic Wall Layers

Figure 56: Magnified view of colonic wall layers and corresponding raw REIMS spectral data

A database was created of spectral datapoints representing colonic mucosa, submucosa and smooth muscle using the REIMS endoscope setup described above. 10 samples of mucosa, 7 of submucosa and 8 of smooth muscle were sampled from 5 patients. The serosal layer was excluded due to technical difficulty in isolation and sampling of this tissue. Analysis of all 3 colonic layers resulted in an overall classification accuracy of 88%. 100% of smooth muscle samples were correctly identified.
5.5.2 Identification of Adenocarcinoma

A database was created of healthy colonic mucosa and invasive adenocarcinoma by sampling each of the different issues using the REIMS endoscope setup. 38 datapoints were generated from analysis of 11 (6 malignant, 5 normal adjacent mucosa) tissue samples from 5 patients. 5 patients provided healthy tissue samples and 5 patients provided tumour samples (patient 4 provided 2 tumour samples from 2 different synchronous tumours in the caecum and rectum). Clinicopathological characteristics are outlined in table 21.
Table 21: Clinicopathological characteristics of tissue samples used in the endoscopic REIMS analysis of colorectal adenocarcinoma and normal adjacent mucosa. BS represents the number of datapoints produced from the REIMS analysis of each sample. This is dependant on the size of the sample with larger samples providing more replicates.

<table>
<thead>
<tr>
<th>Sample number</th>
<th>Macroscopic description</th>
<th>Microscopic description</th>
<th>Location (Surgery)</th>
<th>T stage</th>
<th>N stage</th>
<th>Diff</th>
<th>Neoadjuvant therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 s1 normal BS01</td>
<td>NAM</td>
<td>Healthy mucosa</td>
<td>Rectum (anterior resection)</td>
<td>2</td>
<td>0</td>
<td>Poor</td>
<td>Short course radiotherapy</td>
</tr>
<tr>
<td>1 s2 tumour BS01</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Rectum (anterior resection)</td>
<td>2</td>
<td>0</td>
<td>Poor</td>
<td>Short course radiotherapy</td>
</tr>
<tr>
<td>2 s1 normal BS01, BS02</td>
<td>NAM</td>
<td>Healthy mucosa</td>
<td>Rectum (anterior resection)</td>
<td>3</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>2 s2 tumour BS01, BS02, BS03, BS04</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Rectum (anterior resection)</td>
<td>3</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>3 s1 tumour BS01, BS02, BS03</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Rectosigmoid (high anterior resection)</td>
<td>2</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>3 s2 normal BS01</td>
<td>NAM</td>
<td>Healthy mucosa</td>
<td>Rectosigmoid (high anterior resection)</td>
<td>2</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>4 s1 tumour BS01, BS02, BS03, BS04</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Caecum (right hemicolec tomy)</td>
<td>2</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>4 s2 normal BS01, BS02, BS03, BS04, BS05, BS06</td>
<td>NAM</td>
<td>Healthy mucosa</td>
<td>Caecum (subtotal colectomy)</td>
<td>2</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>4 s3 tumour BS01, BS02, BS03, BS04, BS05, BS06</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Rectum (subtotal colectomy)</td>
<td>3</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>5 s1 tumour BS01, BS02, BS03, BS04, BS05, BS06, BS07</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Rectosigmoid (high anterior resection)</td>
<td>2</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
<tr>
<td>5 s2 normal BS01, BS02, BS03</td>
<td>NAM</td>
<td>Healthy mucosa</td>
<td>Rectosigmoid (high anterior resection)</td>
<td>2</td>
<td>0</td>
<td>Mod</td>
<td>None</td>
</tr>
</tbody>
</table>

Differentiation of healthy invasive adenocarcinoma and normal adjacent mucosa colonic mucosa was possible with an overall classification accuracy of 89.5%, sensitivity 95%, specificity 83% (figure 58). One
A datapoint representing adenocarcinoma was misclassified as normal adjacent mucosa and 3 representing normal adjacent mucosa were misclassified as adenocarcinoma. The misclassified adenocarcinoma sample did not originate from the patient who underwent neoadjuvant radiotherapy.

Figure 58: a) Unsupervised (PCA) and b) supervised (LDA) analysis demonstrating separation of adenocarcinoma and NAM. c) Confusion matrix demonstrating classification accuracy of endoscopic snare analysis of colorectal adenocarcinoma (1) and normal adjacent mucosa (0).

Misclassified samples are shown in table 22.
Table 22: Samples misclassified in iEndoscope analysis

<table>
<thead>
<tr>
<th>Sample Number</th>
<th>Macroscopic Description</th>
<th>Microscopic Description</th>
<th>Clinical Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 s1 BS01</td>
<td>Normal</td>
<td>Healthy mucosa</td>
<td>Anterior resection for rectal poorly differentiated adenocarcinoma. T2N0M0. Short course neoadjuvant radiotherapy.</td>
</tr>
<tr>
<td>4 s2 BS04</td>
<td>Normal</td>
<td>Healthy mucosa</td>
<td>Caecal moderately differentiated adenocarcinoma T2N0M0 and rectal moderately differentiated adenocarcinoma T3, N0, M0. No neoadjuvant therapy</td>
</tr>
<tr>
<td>4 s2 BS06</td>
<td>Normal</td>
<td>Healthy mucosa</td>
<td>Caecal moderately differentiated adenocarcinoma T2N0M0 and rectal moderately differentiated adenocarcinoma T3, N0, M0. No neoadjuvant therapy</td>
</tr>
<tr>
<td>3 s2 BS01</td>
<td>Tumour</td>
<td>Invasive adenocarcinoma</td>
<td>Anterior resection for moderately differentiated rectosigmoid adenocarcinoma. T2N0M0. No neoadjuvant therapy</td>
</tr>
</tbody>
</table>

Using data captured from endoscopic REIMS dissection differentiation between invasive adenocarcinoma and normal adjacent mucosa was possible with a sensitivity of 95% and specificity of 83%. This was a proof of concept experiment to determine whether REIMS technology could be applied to endoscopic tools and as such the data set was small and did not contain any dysplastic tissue.
6. Discussion

Over 800 million pathology tests are performed in the UK alone each year, and over 95% of all clinical patient pathways rely on access to effective pathology services. The ‘gold standard’ for diagnostics is tissue biopsy and histopathological assessment, which is deployed in diagnosis, screening and monitoring of nearly all acute and chronic disease. This is based on the observation of structural features following fixation, embedding, sectioning and staining which is slow, expensive and prone to poor observer-to-observer reproducibility. There is thus an increasing demand for cancer diagnostics with improved quality, safety, efficiency and lower costs that cannot be met by current histological service providers. If cancer outcomes are to be improved, there is also a critical demand for augmented histopathology technologies based on a deeper molecular interrogation of cancer biology to improve real time and intra-operative decision making. Stratified histopathology approaches enabled by multi-omic technologies generate and model rich genotypic and phenotypic data in relation to benchmark patient clinical information. Mass Spectrometry Imaging (MSI) enables the simultaneous measurement of hundreds to thousands of molecular species from intact tissue sections in a spatially resolved manner, and it represents a compelling and transformative technology that exemplifies this challenge. MSI has the potential to deliver a quantum shift in digital pathology services and cancer diagnostics because it augments cellular morphological analysis with highly accurate and robust data on cellular metabolic and proteomic molecular content; critically this analysis is objective and it may be automated. Various imaging mass spectrometric methods have been proposed for the chemical imaging of biological tissues including Matrix assisted Laser Desorption Ionization (MALDI), Secondary Ion Mass Spectrometry (SIMS) and Desorption Electrospray Ionization (DESI). These approaches have allowed fresh tissue samples to be analysed in the ambient environment without subjecting the sample to the high vacuum region of the mass spectrometer. Although these techniques provide novel information on cancer lipid profiling and tumour biology, all imaging MS techniques do still require thin sections of tissue to be prepared and processed and therefore they are not as amendable to in vivo use.
This thesis has demonstrated that REIMS is a potential tool for lipid profiling of human breast and colonic tissue, and has potential future application as an in vivo tool for intra-operative cancer detection. This work describes the development of a novel technique for real time intra-operative tissue diagnostics and in vivo tumour phenotyping through the application of MS analysis of tissue specific membrane lipid species released into the smoke plume during electrosurgical dissection. The system is known as the iKnife and utilizes recent developments in systems biology and spectroscopic technology to provide accurate, real time data on the biology, prognosis and stage of tumours for augmenting clinical decision making in the operating theatre environment. Furthermore, it provides the first evidence for its application in a prospectively collected sample of patients with bowel and breast cancer, where there is a well established unmet need if cancer outcomes are to be radically improved.

Finally, it proposes a methodology for integrated analytical MS based analysis of cancer DE SI-REIMS cross validation. This is essential as it permits spatially resolved biomarkers, identified using MSI, to be rapidly applied to more clinically orientated analysis methods.

6.1 Instrument Development

Preliminary proof of concept REIMS experiments had previously demonstrated that it was possible to generate discriminatory MS lipid profiles from certain tissue types using a modified laboratory based system. The aim of this body of work was to produce an integrated electrosurgery-MS system which was suitable for the clinical environment and also allowed the analysis of tissue types which had previously generated poor quality spectral data, particularly tissue from the GI tract. Systematic development and testing of atmospheric interfaces and MS instrumentation in the analysis of colonic mucosa was undertaken to develop the instrumentation required to meet these goals. A REIMS-iKnife instrumental setup was developed which produced spectral data of adequate quality for the discrimination of tissue types and which was suitable for use within the operating theatre environment. A modified monopolar handpiece suitable for transfer of ionised particles from the point of dissection to the atmospheric interface of the mass spectrometer was produced. A suitable atmospheric interface was developed that incorporated a collision surface to allow further dissociation of large externally
charged aqueous droplets produced by REIMS analysis into multiple smaller ions suitable for MS analysis. To prevent contamination of the instrument and cross contamination of data from different tissue types, a heated coil interface was introduced, which significantly reduced fouling and improved data quality. The result was the Xevo G2-S QTof i-Knife, a fully integrated, heat and noise insulated system, incorporating all elements in a single housing controlled by a touch screen monitor that was suitable for use within the operating theatre, and produced high quality spectral data from the analysis of human colonic samples. An algorithm for lock mass correction was introduced into the data analysis workflow to reduce the effect of mass drift on data interpretation. An automated background subtract function was also developed to improve signal to noise ratio which improves accuracy of data interpretation. At this point in methodology development the background subtract function must be performed manually as it is not yet incorporated into the automated data analysis workflow. Future work will include the integration of this function into the workflow for prospective tissue analysis. The instrument, although now fully integrated, remains large in comparison to other operating theatre equipment and further development work is being undertaken to produce a smaller instrument, based on the same design. Presently a different, laboratory style, atmospheric interface is required for the addition of the IPA solvent into the venturi pump necessary for the analysis of breast tissue. Further interface development is required to produce a fully integrated system, suitable for the analysis of breast tissue within the operating theatre.

6.2 REIMS Analysis of Lipid Metabolism in Cancer Tissue

Multivariate statistical analysis allowed identification of the contribution of individual chemical species to the tissue-specific spectral differences responsible for tissue identification, and thus the identification of candidate biomarkers. The species primarily responsible for the distinction of healthy tissue and invasive cancer in both breast and colorectal samples were phospholipid components of the cell membrane lipid bilayer. As discussed, lipids play important role in a number of critical cellular processes including membrane synthesis, signaling processes and macromolecular synthesis. Although the precise mechanism remains poorly understood, activation of altered lipid metabolism via
upregulation of FASN, de novo lipogenesis, and alterations in membrane phospholipid composition are central to the pathogenesis of many cancers including colorectal adenocarcinoma and invasive breast cancer. The above analyses demonstrate that REIMS analysis of colorectal and breast tissue detects changes in phospholipid profiles between healthy and malignant tissue, which are sufficient for discrimination of each tissue type when multivariate modeling techniques are applied. Lipid species primarily responsible for REIMS distinction of cancer include PCs, PEs, Pls, PS and ceramides, and these findings correlate with other published work. Hilvo et al, as part of The European FP7 METAcancer consortium performed a large integrated gene expression, immunohistochemical and MS based lipidomic analysis of breast cancer and non-malignant breast tissue samples from over 300 patients in an attempt to characterize the molecular lipid changes in breast cancer. Over 600 metabolites were structurally identified including complex lipids, with phospholipids and particularly phosphocholines, being the most consistently altered species in cancer tissue. They too found that membrane phospholipids including PCs PEs, and Pls, as well as sphingomyelins and ceramides were the most altered in breast cancer whilst triglycerides were largely unaltered. Guo et al, using MALDI MSI, reported a significant increase in monounsaturated phosphocholines (Inc. PC(32:1), PC(34:1) and PC(36:1) in 6 different solid organ malignancies (breast, colorectal, lung, oesophageal, gastric and thyroid) compared to adjacent normal tissue and a significant increase in polyunsaturated PCs (Inc. PC (38:4)) and phosphatidic acid (inc PA (40:5) in adjacent healthy tissue compared to cancer tissue. Eighteen common variables were found to be strongly associated, positively or negatively, with all 6 cancer types including breast and colorectal cancer. These included 10 phospholipids with 3, PC(32:1), PC(34:1) and PC(36:1), all found to be increased in cancer (p<0.01) and 7, PC(38:4), PC(38:6), PA(36:2), PA(38:3), PA(40:5),PE(38:4), PI(38:4), all found to be increased in normal adjacent tissue. (p<0.01). There were consistencies in the findings of this study and our own with 5 of the same phospholipids species also found to be significantly increased or decreased in the REIMS analysis of breast cancer tissue (PC(32:1), PC(34:1), PA(40:5), PE(38:4), PI(38:4)). Although no identical lipid species were found to be significant in the REIMS detection of colorectal cancer, the trend for increased levels of phosphatidic acid in healthy...
tissue compared to cancer was maintained and the same species (PE, PI, PA, PC) were primarily responsible for the distinction of the two tissue types.

6.3 REIMS Analysis of Colorectal Tissue

Normal, benign and malignant colorectal tissue was sampled from 40 patients and analysed ex vivo by handheld monopolar REIMS using the instrument and data processing methodology developed in the early stage of this work. A prospective database was built which allowed the interrogation of the dataset to answer questions of clinical importance.

Differentiation of colorectal adenocarcinoma from normal adjacent mucosa was possible, based on multivariate analysis of MS lipid profiles of each tissue type as measured by REIMS, with a sensitivity of 93.6% and specificity of 93.8%. The positive and negative predictive value of the model were both high, with a positive predictive value of 93.6% and negative predictive value 93.8%. As a diagnostic test for the presence of tumour at the resection margin, the high positive predictive value is important as it represents a low probability of cancer cells being left in situ when the algorithm predicts a negative result.

The REIMS system was able to diagnose dysplasia from healthy mucosa with a sensitivity of 91% and specificity 93%, however distinction between dysplastic and malignant tissue was less accurate with 78% of the small cohort of dysplastic samples classified as malignant. There are clear, detectable, lipid alterations between normal healthy mucosa and invasive cancer, however, the vast majority of lipid changes significant in the distinction of invasive cancer and normal mucosa were also detected in the differentiation of healthy mucosa and dysplasia. Dysplasia, derived from the Greek “malformation” or “morphologic structure different from what normally is expected, is a spectrum of disease from low grade to high grade. High grade dysplasia can be difficult to distinguish from invasive cancer, which requires morphological evidence of invasion in the form of disruption of the muscularis mucosae by neoplastic cells and the presence of desmoplasia. This is particularly challenging in the examination of superficial or poorly orientated biopsy specimens and, clinically, high grade dysplasia is often managed
as invasive cancer. Indeed invasive cancer is frequently found in specimens from patients operated on for high-grade dysplasia.

Of the 13 most significant species identified in the differentiation of invasive cancer and normal adjacent mucosa, 9 were also significantly under or overexpressed between dysplasia and normal adjacent mucosa. Three of the remaining species, Prostaglandin D2, PE(-32:1) and PS(40:1), were not significant and PE (34:0) was found to be significantly higher in normal tissue than in dysplasia, despite being increased in invasive cancer. Only Prostaglandin D2 and PE (34:0) were significant in the differentiation of dysplasia and malignancy. This suggests that the majority of these changes in lipid metabolism may occur in the transition from normal to dysplastic tissue rather than at the point of transition from dysplasia to invasive malignancy. This correlates with the fact that it is possible to differentiate between NAM and dysplasia and NAM and invasive cancer, using REIMS analysis, with a high level of accuracy, but achieves poor results in the differentiation of dysplasia and invasive malignancy. The biochemical changes which produce the morphological similarities between high grade dysplasia and malignancy may lead to similarities in the spectral profiles leading to challenges in REIMS based differentiation. Normal healthy mucosa could be distinguished from abnormal (dysplastic and malignant tissue with a sensitivity of 91% and specificity of 93% providing further evidence that there are significant detectable biochemical and lipid alterations in the process of dysplastic change.
Sample size in the dysplasia subgroup was very small and contained 3 samples of low grade dysplasia sampled from large adenomas and one sample of high grade dysplasia sampled from a resection specimen containing invasive cancer. Further sampling and analysis is required to investigate whether lipid changes can be detected between low and high grade dysplasia and to determine, with more equal sample groups whether REIMS analysis can distinguish between dysplasia and invasive malignancy. The biochemical changes which produce the morphological similarities between high grade dysplasia and malignancy may lead to similarities in the spectral profiles leading to challenges in REIMS based differentiation.

Well established histological biomarkers are used in routine clinical practice for prognostication or for stratifying surgical or adjuvant therapy. However, these data are generally provided in the post-operative setting. It is rare, particularly in early rectal cancers, for pre-operative histology to guide the
surgical, or neoadjuvant strategy. Improved data on tumour biology, stage, or local or systemic spread would be of significant importance in the planning phase.

Rectal and colonic cancers are different clinical entities, each displaying characteristic features, attributable to differing embryological development, physiological circumstances and environmental factors. These two cancers, for many years considered a single disease, also display individual therapeutic and prognostic features. It is proposed that proximal and distal colonic cancers arise through different pathogenetic mechanisms\textsuperscript{220,221} with proximal tumours more likely to be related to the nucleotide instability pathway, as microsatellite instability\textsuperscript{50} and distal cancers associated with specific chromosomal instability, in particular p53 mutations\textsuperscript{222}. Recently site of origin from the midgut, rather than hindgut, has been proposed as an explanation for gene hypermethylation seen more commonly in right sided tumours\textsuperscript{223}. Using REIMS analysis, it was possible to distinguish between colorectal adenocarcinoma originating in the rectum/rectosigmoid junction from those originating in the remainder of the colon with an overall accuracy of 87%. There was only one sample of colonic cancer originating distal to the splenic flexure but proximal to the rectosigmoid in this study and therefore separation based on colonic location was not possible. Due to the differing lipid signatures produced during REIMS analysis of colon and rectal tissue, future REIMS studies of colorectal tissue should treat colon and rectal adenocarcinoma as distinct disease processes.

MRI is the current gold standard modality for determining T stage, EMVI, lymphatic involvement, and CRM involvement in rectal cancer, however sensitivity and specificity for the prediction of these features are reported at 87% and 75% respectively for T stage and 77% and 71% for lymphnode involvement when compared to postoperative histology\textsuperscript{55}. In samples of rectal cancer, using ex vivo REIMS analysis, it was possible to differentiate between T2 and T3/4 lesions with a sensitivity and specific of 92% and 67% (AUC 0.87) and to identify tumours with lymphnode involvement (N0 vs N1/2) with a sensitivity and specificity of 70% and 88% (AUC 0.93%). EMVI, defined as tumour cells actively invading the veins beyond the muscularis propria is a strong prognostic biomarker of poor prognosis\textsuperscript{224}, and its presence allows patients to be substratified into higher and lower risk subgroups. The KRAS
gene mutation, most commonly localised to codons 12 or 13, is present in approximately 30-50% of colorectal cancers and is a predictor of very poor response to anti EGFR chemotherapeutic agents. In addition, the presence of KRAS mutation is independently associated with increased tumour aggressiveness and significantly poorer survival after diagnosis \(^{225,226}\). Currently the presence of EMVI is detected on pre-operative MRI and the presence of KRAS mutation resected tumour or biopsy tissue first examined by a pathologist to evaluate the tumour content of the sample. Microdissection of the tissue is performed before DNA is extracted and mutation analysis is carried out to determine the KRAS status. Current methods for detection of KRAS mutation in the clinical environment include strip assays, pyrosequencing, high resolution melt analysis and MALDI ToF MS analysis of RNA. All require DNA extraction, amplification and mutation analysis to determine KRAS status. The limit of detection (the lowest amount of mutant DNA in a background of wild-type DNA at which a mutant sample will produce a positive result 95% of the time) range from 0.77 (PCR assay) – 10% (MALDI-ToF). Ex vivo REIMS analysis of colonic and rectal adenocarcinoma demonstrated a high level of accuracy in the identification of KRAS mutation (sensitivity 85%, specificity 82%. AUC 0.95) and the presence of EMVI (sensitivity 77%, specificity 89%. AUC 0.93). Formal results were acquired from the final histopathology report of resected cancers. Level of mutant DNA was not reported and as such no preliminary investigation into the level of mutation required for appositive result could be undertaken.

Detection of T stage and lymphnode involvement was less accurate in analysis of colonic lesions compared to tumours arising in the rectum. Further analyses are required to investigate the lipidomic differences between colonic and rectal cancer and the reasons why distinction of some individual clinical features was more challenging in colonic cancer, particularly taking into account the proposed underlying pathogenetic differences between right and left sided colonic lesions.

6.4 REIMS Analysis of Breast Cancer

An instrumental setup was developed and tested which is suitable for the REIMS analysis of lipid rich breast tissue. REIMS with IPA venturi produces spectral data from the analysis of both fibrous tumour
tissue and fibrofatty/glandular healthy breast tissue with sufficient discriminatory phospholipid signal for the identification of each tissue type. A prospective database of healthy, benign and malignant breast tissue was built using this system. Differentiation of invasive breast cancer and normal healthy breast tissue was possible with a sensitivity of 87.1%, specificity 89.1% and AUC 0.95 based on multivariate analysis of MS lipid profiles of each tissue type as measured by REIMS. When combined in a prospective database with healthy tissue and invasive cancer, the predictive power of REIMS for the identification of benign lesions was low. Only 31.6% of fibroadenoma were correctly identified with a large percentage (47.4%) misclassified as healthy breast tissue. However, when directly compared with healthy tissue, 79% of fibroadenoma and 85% of healthy tissue were correctly identified. Unsurprisingly, the most significant alterations in lipid species were detected between healthy tissue and invasive cancer with the lipid profiles of fibroadenoma and healthy tissue being more similar. Most lipid species significantly altered in cancer were not significant in the distinction of healthy tissue and fibroadenoma with the exception of 2 PA(36:0) and PG(38:4) which were found to be increased in both cancer and fibroadenoma compared to healthy tissue. Despite the misclassification of benign samples as healthy when included in a dataset with all 3 tissue types, when directly comparing healthy tissue and fibroadenoma it was possible to identify fibroadenoma with a sensitivity of 79% and specificity of 85%. The number of benign tissue samples was small and, although results suggest that REIMS may have potential has a margin detection tool in breast cancer resections, further studies are required to determine the adequacy of REIMS as a tool for the differentiation of multiple histological tissue types arising from the human breast for diagnostic purposes. Unfortunately this dataset did not contain any samples of breast carcinoma in situ. Tissue samples were harvested from patients with DCIS and LCIS with the intention of including these into the database, but on histological validation the samples were all found to be invasive disease or healthy breast tissue. The identification of in situ malignancy is a key component of any in vivo diagnostic tool due to the high incidence of in situ disease at resection margins and the impact that this has on re-excision rates. This was a proof of concept study to develop a suitable instrumental setup for REIMS analysis of lipid rich breast tissue and to determine whether REIMS analysis is a suitable tool for the analysis of human
breast tissue. Future work will focus on identification of in situ malignancy and discrimination between this and healthy tissue and invasive disease.

It has been proposed, in previously published work, that phospholipids may provide future targets for therapeutic modulation therapeutic and may also have diagnostic and prognostic potential in breast cancer\textsuperscript{227}. Increased expression of FASN and the downstream effects on altered membrane phospholipid composition have been associated with poor prognosis\textsuperscript{105}, aggressive tumour phenotypes\textsuperscript{110} and increased risk of recurrence\textsuperscript{105} in breast cancer. Glycerol-3-phosphate acyltransferase (GPAM), a key enzyme in the biosynthesis of triglycerides and phospholipids has been shown to be differentially expressed in malignant breast tumours. GPAM expression is more strongly associated with increased levels of phospholipids, especially phosphatidylcholines in breast cancer cells than triglycerides\textsuperscript{228} and also correlates with clinico-pathological characteristics, including hormone receptor status and overall survival\textsuperscript{228}. It is therefore logical that the downstream alterations in membrane phospholipid composition associated with alterations in GPAM expression may also be associated with certain tumour characteristics. Although the largest observed variation in lipid profiles have been reported between healthy breast tissue and tumour tissue, tumours of differing grade, hormone receptor status and those with reduced overall survival have been identified based on their phospholipid profiles\textsuperscript{229}. MAS NMR based studies have also demonstrated associations between the metabolite profiles produced by analysis of breast tumour tissue and prognosis\textsuperscript{157,230,130,231}, lymph node involvement\textsuperscript{129,127} and tumour grade\textsuperscript{232,125}. REIMS spectral profiles of tumour tissue demonstrated specificity for some individual tumour biological characteristics. Tumour type and HER 2 receptor status accounted for the greatest difference in metabolic profiles of breast tumour tissue, and could be identified with classification accuracies of 92.9% and 85.7% respectively. ER status was less significant with a classification accuracy of 72.9%. This does not correlate with findings reported in the recent paper published by Hilvo et al who found that ER profiling was the main discriminator of molecular signatures in breast cancer tissue\textsuperscript{233}. Previous transcriptomic analysis has demonstrated an association between HER2 signaling and lipid metabolism\textsuperscript{234} and HER2 overexpression has been shown to confer a
lipogenic and more aggressive phenotype\textsuperscript{110}, alterations proposed to be secondary to increased translation of FASN and a resulting increase in HER2 signalling\textsuperscript{114}. HER2 status was a strong discriminating factor in REIMS analysis of breast cancer tissue. This classification could not be explained by the presence of other discriminating clinical factors between the HER2 positive and negative groups as HER2 positive tumours were of multiple grade (G2 and G3), had differing ER status and one patient had neoadjuvant chemotherapy. The same study, which reported the high significance of ER status in the alteration of lipid profiles of breast cancer, also found that after ER status, tumour grade affects the lipid profile of breast cancer tissue most radically. The classification accuracy for the REIMS based discrimination of grade 1 to grade 3 breast cancers was low at 72.9\%, however, REIMS analysis in addition to several other profiling techniques, demonstrated distinct biochemical differences between grade 1 and grade 3 disease. No grade 1 samples were misclassified as grade 3 or vice versa, instead a gradual biochemical transition was visualized from low grade to high grade disease. Histological tumour grade is entirely dependent upon subjective examination of intracellular morphology with an inherent risk of inter-rater variability and based on a seemingly arbitrary classification system. This does raise the question of the relevance of tumour grade in a future molecular era of precision medicine and how this subjective algorithm can be applied to future classification systems based on objective biochemical data. However, since the initial demonstration in 1991\textsuperscript{235} of the significant prognostic value of tumour grade in breast cancer using the Nottingham Grading System, confirmation in multiple studies\textsuperscript{236,237,238} has lead to the inclusion and significant weighting (at least equal to lymphnode status) of tumour grade in prognostic indices and also in algorithms used for determining need for adjuvant therapy. Tumour grade is an independent prognostic indicator in certain tumour types, in fact, those with histological grade 1, stage II cancer have equal survival to those with histological grade 3 stage I disease\textsuperscript{239}, thus the inherent biological features leading to the morphological changes which determine histological grade are important in determining tumor behavior. It is not to be expected that tumour grade, as a continuous, subjectively assigned, biological variable, should produce the same distinct classes as biochemical profiling data, however, profiling studies (genomic, transcriptomic and immunohistochemical) have demonstrated distinct gene expression and molecular profiles produced by
tumours of different histological grades. This has drawn the conclusion that perhaps low grade tumours do not progress to high grade tumours, but that they are distinct pathologies with differing tumour biology and that tumours classified histologically as grade 2 simply do not belong to either of the 2 identified distinct cohorts. One gene expression profiling study found that 97 genes, predominantly associated with cell cycle regulation were associated with ER positive tumours with histological grade 1 or 3\textsuperscript{240}, but no specific gene expression profiles for histological grade 2 cancers could be identified. The previously described MS lipid profiling study published by Hilvo et al successfully differentiated between grade 1/2 and grade 3 cancers and similarly, MAS NMR studies successfully differentiated between “high” and “low” grade cancers based on their metabolic profiles. Although no one has successfully produced 3 distinct classes in complete accordance with histological grade, based on biochemical profiles, there is evidence that the morphological change associated with decreased levels of differentiation and tumour aggressiveness are associated with distinct biochemical changes, demonstrated in the differentiation of high and low grade tumours by their distinct lipid profiles.

6.5 Identification of Positive or “Close” Margins

The spatial resolution of REIMS is limited to the accuracy of surgical dissection. Studies using high resolution MS profiling techniques have demonstrated the presence of peritumoural metabolic changes\textsuperscript{141,151}, consistent with the theory of field cancerisation\textsuperscript{241}, i.e. that cells surrounding primary tumors have been exposed to similar carcinogenic effects, thus these histologically healthy cells already exhibit metabolic changes associated with cancer progression. Using these techniques it was possible to distinguish between tumour adjacent and tumour remote tissue and describe the observed biochemical biochemical differences between them\textsuperscript{151}. After many years of controversy and regional variation on what constitutes a “positive” or “close” margin in breast cancer surgery, guidance produced by the Society of Surgical Oncology and the American Society for Radiation Oncology in 2014\textsuperscript{242} suggests that the definition of a positive margin in breast conserving surgery for cancer should be “tumour on ink” or tumour cells in direct contact with the inked margin. Based on a meta-analysis of 33 studies including more than 28,000 women, the recommendation also confirms that whilst a positive resection margin
doubles the risk of ipsilateral breast cancer recurrence, clear margins, as defined by no ink on tumour offer the lowest risk of recurrence and wider margins do not reduce this any further. However, in the UK, the most recent consensus statement from the Association of Breast Surgeons (Bournemouth, UK, June 2015) suggests that if tumour cells are present within 1mm of the inked margin, then re-excision should still be considered. A system for in vivo margin detection would ideally alert the operator to the presence of tumour cells in very close proximity to the dissection line, to allow widening of the plane of dissection to prevent cutting through tumour tissue. However, considering that wider resection margins of 1-2mm are no longer considered necessary, a system which alerts the operator to dissection through tumour cells would allow a small volume of additional tissue to be removed at the positive margin until the system no longer detects the presence of tumour cells. It remains unclear from this provisional work whether REIMS analysis detects a change in metabolic signature in peritumoural tissue or at what distance from tumour cells this may occur. This dataset contained several breast cancer tissue samples in which the sampling point contained both dense tumour cells and adipose tissue or tumour cells were presence <2mm away. Subset multivariate analysis demonstrated that where the sampling point included dense cancer cells, this sample did not cluster with samples of healthy breast tissue but occupied a space on the scores plot that was quite distinct. This sample was classified by the algorithm as cancer, and this would be necessary in any margin detection system, as cutting directly through the border of cancer and healthy tissue is would lead to a positive margin as defined by recent guidelines, and this must be detected. Samples containing tumour cells in close proximity to the sampling point occupied similar space to healthy tissue on the scores plot. Where tumour cells were >1mm from the sampling point, the sample was classified by REIMS as healthy breast tissue. Similarly, it would be necessary for a margin detection system, in the current era of “tumour on ink” to classify this as a negative margin to avoid unnecessarily high volume of excised tissue. Two samples contained sparse tumour cells 0.1mm from the sampling point. One of these samples was classified as cancer and one as healthy tissue. REIMS is a destructive technique, and has a low spatial resolution, which increases potential for error in reporting the presence of a small number of tumour cells within the sample after analysis, particularly in the detection of single scattered tumour cells within the line of dissection. It is
very difficult to achieve more accurate spatial resolution using handheld monopolar REIMS due to both the limits of human accuracy and minimal size of the sampling tip (0.5mm) required for adequate tissue dissection and the generation of sufficient signal. For the purposes of future experimental work and database construction, automated REIMS sampling should be developed, to sample each piece of tissue in a raster pattern with predetermined distances between each sampling point. This would allow more accurate documentation of chemical profiles produced during dissection through cancer, the tumour microenvironment and healthy tissue, to identify whether peritumoural metabolic changes can be detected by REIMS and at what distance from tumour cells. More accurate spatial resolution is possible using alternative, non-destructive mass spectrometric imaging techniques such as MALDI and DESI MSI. Highly spatially resolved spectral databases could be built using alternative MS imaging platforms and utilised in the used for the classification of REIMS samples where there is any ambiguity about the nature of the sample. In addition, a significant rate-limiting step in the implementation of REIMS technology is construction of the very large, mass spectral databases of each tissue type required prior to tissue identification. Mass spectrometry imaging provides high volume, spatially resolved data from the analysis of human tissue samples but is not suitable for in vivo sampling. Using methods to reduce the effects of different MS platforms, DESI MSI could be utilised to build the large, highly accurate lipid profiling databases required for in vivo REIMS analysis.

6.5 Cross-Platform MS Analysis

We aimed to build a system which would allow such cross-platform utilisation of data. Firstly a prospective DESI MSI databases of colorectal tissue was built and proved sufficient for the differentiation of colorectal adenocarcinoma and healthy mucosa with an overall classification accuracy of 93.8%, sensitivity 91% and specificity of 96%, a level of accuracy comparable to REIMS analysis. As with REIMS analysis, the model performed poorly in the discrimination of dysplasia with no spectra representing dysplasia classified correctly and 95.2% of spectra misclassified as healthy mucosa. Similar lipid species were responsible for the identification of adenocarcinoma using both MS methods. Seven lipids, PE(-32:1), PI(38:3), PI(38:2), PI(38:4) PI(40:5), PG(36:4) PS(40:1) were statistically significant in the
differentiation of cancer and normal adjacent mucosa using both REIMS and DESI analysis. Using this data, a cross platform data handling method was developed by Mr Ottmar Golf and Dr Kirill Veselkov (Division of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London) to reduce variation in the data secondary to ionization method and MS instrument used. Classification of tissue types using cross platform normalized chemical profiles lead to the identification of colorectal adenocarcinoma and normal adjacent mucosa sampled by REIMS utilizing a database built using DESI MSI data with an overall accuracy of 84%. This is significantly lower than the 93% classification accuracy of REIMS in the identification of colorectal malignancy, in particular, specificity reduced from 94% to 75% using the cross platform methodology. However, this was a preliminary study using an experimental method for the utilization of a training set of data built on one MS platform for the classification of data sampled on another. Further testing of the system with higher sample numbers and ongoing system developments are in progress and will determine whether this will be a feasible method for more efficient building of REIMS databases in the future.

6.6 iEndoscope

Due to the increasing demand on endoscopy services throughout the UK, and increasing use of minimally invasive techniques in the management of luminal GI tract pathology, there is clinical need for a tool which could provide tissue specific information during endoscopic examination and intervention in order to differentiate between dysplastic and hyperplastic polyps in vivo, identify the presence of residual disease at the resection margin and determine depth of tissue dissection to improve accuracy and safety of endoscopic resection. It was possible to create an endoscopic snare suitable for the transfer of gas phase ionic species from the point of colorectal tissue dissection into the mass spectrometer, which produced sufficient quality spectral data for the differentiation of adenocarcinoma and healthy colonic mucosa (classification accuracy 89.5%) and individual colonic wall layers (classification accuracy 88%). If this system is effective in vivo, it could provide objective, near real time identification of malignant change in the colon and rectum and improve the effectiveness and
safety of endoscopic intervention for colorectal neoplastic lesions. It could also circumvent the need for histological examination of all excised tissue as tissue with no malignant potential e.g. hyperplastic polyps could be left in situ or removed and discarded (the so called diagnose and leave behind or diagnose and discard process). Near real time identification of colonic wall layers could also alert the endoscopist to the risk of impending perforation during complex endoscopic resections of large lesions. Unfortunately this small proof of principle dataset did contain any samples of dysplasia and therefore application in the differentiation of dysplasia and malignancy and dysplastic and hyperplastic polyps cannot be conferred. However, results from ex vivo colorectal REIMS experimental work (chapter 5.2) demonstrate differentiation of dysplastic tissue from normal adjacent mucosa with an overall accuracy of 96.8% using handheld monopolar REIMS. This suggests that application of the endoscopic snare system to the differentiation of hyperplastic (normal mucosa) and dysplastic polyps is feasible. Preliminary results of the monopolar REIMS analysis also demonstrated differences in lipid profile associated with T and N stage, KRAS and EMVI status. Application of this to an endoscopic tool has potential to improve decision making around the management of early colorectal cancer (T1, N0 tumours) or malignant polyps. As many as 10-15% of patients undergoing local resection may have undetected lymph-node positive disease and would benefit from formal surgical resection. A method for detecting this, by direct in vivo analysis of the tumour tissue itself, rather than on post operative histological examination after the surgical decision making process is complete, has potential to improve the clinical management of these patients. Current techniques for in vivo, endoscopic tissue diagnostics have shown an acceptable sensitivity and specificity in the differentiation of adenomas from hyperplastic polyps and could make useful adjuncts in the diagnose and leave behind or diagnose and discard process. However, image based techniques are user dependent and highly subjective, have a learning curve and can be time consuming. Some require the use of local or systemic contrast/fluorescence agents, which risk hypersensitivity and can lead to lengthy procedures. Due to the high throughput in the endoscopy department, on table diagnostic tools must be rapid, accurate, and cost effective. There have been advances in the use of radiofrequency spectroscopy, elastic scattering spectroscopy and Raman spectroscopy to provide objective, endoscopic
histological information, however none of these technologies have, as yet, combined satisfactory accuracy with a system which is suitable for use in the clinical environment. Advantages of a REIMS based tool over alternative optical tools currently in development are its objective, user independent nature. REIMS requires very little additional training; results are displayed on-screen requiring no further interpretation, the technology is incorporated into standard tools used for endoscopic tissue sampling and dissection and does not require the use of any endogenous contrast or fluorescence agents.

This was a feasibility study and was limited by small sample size and lack of dysplastic issue. Further analysis to determine the accuracy of the system in the detection of tumour characteristics was not carried out due to the small sample size. Although samples were analysed using a modified snare within a colonoscope, the ex vivo nature of this experiment did not take into account confounding factors which are present during in vivo colonoscopy including the presence of faecal material, the use of insufflation and suction and the use of electrosurgery and production of surgical smoke within a confined space. Larger ex vivo studies with the inclusion of dysplastic tissue and in vivo studies to determine quality of data captured in the clinical environment are required to validate these results and to determine whether the REIMS capable endoscope will prove to be a useful tool for real time endoscopic assessment of tissues in the future.

6.7 Limitations

The purpose of this body of work was primarily the development and feasibility testing of an electrosurgery-MS based tool for near real time human tissue characterization. As such, sample numbers were low, particularly for tissue types which undergo excision less commonly such as benign lesions and tissue types which proved difficult to sample due to risk of disrupting post operative histological analysis, such as in situ cancer. This led to difficulty in full subset analysis with equal numbers of tissue samples and as such, analysis of some individual clinical variables was not possible. The effect of neoadjuvant therapy on REIMS analysis could not be determined here as the clinical response to therapy often reduced tumour size so that it was not safely possible to take research
samples and as such numbers were small. This preliminary work was carried out *ex-vivo* without the presence of environmental factors which may influence quality of spectral data obtained, particularly in endoscopic analysis where the environmental conditions of the gut may significantly influence data captured. Due to the short timescale, it was not possible to provide clinical outcome data for patients recruited into this study. Although the final MS platform (Waters Xevo G2-S QTof i-Knife) was designed and tested for compatibility with the operating theatre environment, no feedback on instrument design or usability has been sought from clinicians at this stage.
Summary Conclusions and Future Work

Recent advances in biomolecular science and technology have lead to rapid growth in the field of –omics technologies resulting in increased use of top down systems biology approaches in cancer research. Metabolic profiling in health and disease is a rapidly developing field; underpinned by the philosophy that metabolic phenotyping, when applied in the clinical environment, could improve decision making before, during and after surgery by providing diagnostic and prognostic information that is more accurate and personalized than current biochemical approaches. The primary aim of this body of work was the development and testing of a tool for future near real time, intra-operative tissue identification by analysis of cellular lipid profiling information contained within the byproduct of electrosurgical dissection. A direct combination of a dissection tool and a tissue identification system, the iKnife could represent an effect method for surgical margin control as it is rapid and objective, has no learning curve, utilizes the byproducts of one of the most commonly used operating tools, requires no tissue pre-processing or exogenous agents and can be applied to any tissue or pathology. The entire sampling process taking less than 3 seconds compared to over 30 minutes for frozen section analysis. The presented data demonstrates development of a functioning REIMS system, which is applicable to the analysis of colorectal and breast tissue as a tool for cancer detection. Preliminary work was also carried out into the REIMS distinction of tumour prognostic factors. Promising results were demonstrated in the accurate detection of the presence of KRAS mutation in colorectal cancer and HER2 receptor status in breast cancer. There is growing recognition that some tumors may be biologically more favorable than others and that the stage of a tumour at diagnosis is not the only factor to be considered when predicting survival outcomes. The number of variables evaluated to formulate treatment plans have multiplied as our understanding of tumour biology improves and the trend towards more personalized therapies increases. Future treatment options are more likely to depend on such individualized features. Further understanding of the diagnostic and prognostic value in the lipid alterations observed by REIMS associated with individual tumour characteristics could lead to improved prognostic and therapeutic targets, not simply by the pre-operative identification of
established prognostic indicators such as T and N stage but the identification of new lipidomics based
prognostic indicators based on lipid profiling of cancer tissue itself.

An early, universal feature of malignancy is the activation of de novo lipogenesis and development of a
lipogenic phenotype\textsuperscript{102}. However, the precise molecular lipid changes associated with malignant
transformation remain poorly understood. REIMS analysis provides an opportunity to study the end
products of this lipogenic change, the cellular lipids, at the molecular level to improve our
understanding of the processes which drive tumour growth and development. The results presented
here have identified candidate lipid markers altered in the transition from healthy tissue to invasive
breast and colorectal cancer, several of which correspond with those identified by other authors.

Future work will utilise the instrument and analytical platforms developed here to further investigate
the future potential of REIMS as a much-needed tool for intra-operative tissue diagnostics. Large scale
ex vivo databases will be built including adequate and equal numbers of healthy tissue, dysplasia and
malignancy to determine whether dysplasia and benign lesions can be differentiated from both healthy
and malignant tissue simultaneously. These databases should be very large to allow each clinical
characteristic to be managed as an independent variable to allow the effect of each on the lipid profile
of cancer tissue to be investigated. Certain clinical features account for large amount of variance in the
dataset and therefore outweigh the smaller effects of others, and therefore each variable must be
isolated to test the discriminatory power of the system for the identification of each. Further work will
be carried out to evaluate the influence of the presence of single tumour cells on the designation of a
tissue sample. Clinically, the presence of a single tumour cell at the resection margin constitutes a
positive margin, however, in the REIMS algorithm normal tissue is not a negative or absence of data
consistent with the presence of cancer cells. Analysis of normal tissue also yields a positive result. This
poses the question of how to ensure a single cancer cell is detected in a sample of normal tissue, as
both have equal weighting in data analysis. It will be necessary, in the first instance, to build databases
containing such samples, accurately classified histologically as cancer to produce a model which
recognizes that when the positive result of normal tissue and the positive result of cancer occur in the
same sample, the sample must be classified as malignant. Accurate detection of single or scattered tumour cells within histologically normal tissue is essential if REIMS is to be considered as a surgical margin assessment device in the future.

*In vivo* studies are required to determine whether REIMS analysis can deliver real-time histological analysis, prospectively, in the clinical environment with accuracy comparable to current pathological analytical standards. In the final stages this will require prospective clinical trials comparing REIMS intra-operative tissue analysis with the current gold standard used in healthcare within the UK, which remains the histological frozen section. *In vivo* endoscopic REIMS studies are required to assess the impact of factors such as the presence of faecal material, insufflation and suction on the quality of the spectral data produced. Currently the only surgical device, which has been developed and tested for REIMS use, is the handheld monopolar electrosurgery handpiece for open surgical dissection. We plan to develop a device which is compatible with laparoscopic surgery, based on the same principles of transmission of the ion rich smoke plume produced during laparoscopic electrosurgical dissection. More recently, ergonomic and efficient modern dissection and vessel sealing tools, such as the Harmonic Scalpel™ (Ethicon Endosurgery), LigaSure™ (Covidien Medtronic) and Thunderbeat™ (Olympus), are being utilized to allow more complex operations to be performed laparoscopically. Such devices are the preference of many surgeons over simple electrosurgery tools when carrying out complex laparoscopic work. The Harmonic Scalpel™ cuts and coagulates tissue simultaneously using ultrasonically generated frictional heat energy, the LigaSure™ utilises electrically generated bipolar energy and the Thunderbeat™, a combination of both. All of these devices produce surgical smoke as a by-product of tissue degradation. No experiments have been carried out to determine whether the same lipid species can be identified by REIMS analysis of surgical smoke produced by any of these devices, as are seen during REIMS analysis of smoke produced by electrosurgical devices. It appears likely, as these lipid species are released at the point of thermal degradation of tissues, that the use of any device which generates heat energy, whether by bipolar electrically generated or ultrasonically generated heat energy should result in the release of cell membrane constituents. However, future
experimental work must be carried out to determine whether modified, REIMS capable versions of such devices could be effective in generating spectral data of sufficient quality for analysis of human tissue during laparoscopic surgery.

If *in vivo* endoscopic studies demonstrate the production of spectral data of sufficient quality for tissue discrimination then a tool for REIMS analysis during transanal endoscopic microsurgery will be designed and tested. A qualitative analysis of user experience will be carried out to optimize instrument settings and data visualization system for use by clinical staff during in vivo testing.

Further development of the MS cross platform workflow will also be carried out to improve the accuracy of the system for cross platform tissue sampling and database building to optimize the efficiency of large scale database construction.

The lipidomic changes associated with oncogenesis and cancer progression, as measured by REIMS, will be systematically investigated. Lipid changes associated with tumour characteristics and clinical outcomes will be monitored. The information provided by this downstream lipidomic analysis, when combined with findings from genomic and proteomic analysis, could provide valuable insight into the complex intracellular changes associated with cancer growth and development.
8 References


85. Hormone Receptor Status and Breast Cancer. at <http://www.breastcancer.org/symptoms/diagnosis/hormone_status>


137. Moreno, A., Rey, M., Montane, J. M., Alonso, J. & Arús, C. 1H NMR spectroscopy of colon tumors and normal mucosal biopsies; elevated taurine levels and reduced polyethyleneglycol absorption in tumors may have diagnostic significance. *NMR Biomed.* **6**, 111–8


## Appendix 1: Colorectal Cancer TNM Staging System (7th ed.)

<table>
<thead>
<tr>
<th>Primary tumor (T)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TX</td>
<td>Primary tumor cannot be assessed</td>
</tr>
<tr>
<td>T0</td>
<td>No evidence of primary tumor</td>
</tr>
<tr>
<td>Tis</td>
<td>Carcinoma in situ: intraepithelial or invasion of lamina propria</td>
</tr>
<tr>
<td>T1</td>
<td>Tumor invades submucosa</td>
</tr>
<tr>
<td>T2</td>
<td>Tumor invades muscularis propria</td>
</tr>
<tr>
<td>T3</td>
<td>Tumor invades through the muscularis propria into the pericolorectal tissues</td>
</tr>
<tr>
<td>T4a</td>
<td>Tumor penetrates to the surface of the visceral peritoneum</td>
</tr>
<tr>
<td>T4b</td>
<td>Tumor directly invades or is adherent to other organs or structures</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regional Lymph Nodes (N)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NX</td>
<td>Regional lymph nodes cannot be assessed</td>
</tr>
<tr>
<td>N0</td>
<td>No regional lymph node metastases</td>
</tr>
<tr>
<td>N1</td>
<td>Metastasis in 1-3 regional lymph nodes</td>
</tr>
<tr>
<td>N1a</td>
<td>Metastasis in 1 regional lymph node</td>
</tr>
<tr>
<td>N1b</td>
<td>Metastasis in 2-3 regional lymph nodes</td>
</tr>
<tr>
<td>N1C</td>
<td>Tumor deposits in the subserosa, mesentery, or nonperitonealized pericolic or perirectal tissues without regional nodal metastasis</td>
</tr>
<tr>
<td>N2</td>
<td>Metastasis in 4 or more regional lymph nodes</td>
</tr>
<tr>
<td>N2a</td>
<td>Metastases in 4-6 regional lymph nodes</td>
</tr>
<tr>
<td>N2b</td>
<td>Metastases in 7 or more regional lymph nodes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distant Metastases (M)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0</td>
<td>No distant metastases</td>
</tr>
<tr>
<td>M1</td>
<td>Distant metastases</td>
</tr>
<tr>
<td>M1a</td>
<td>Metastasis conformed to one organ or site</td>
</tr>
<tr>
<td>M1b</td>
<td>Metastases in more than one organ or site or in the peritoneum</td>
</tr>
</tbody>
</table>
### Appendix 2: Colorectal cancer tumour stage as determined by TNM classification groupings

<table>
<thead>
<tr>
<th>Stage</th>
<th>T</th>
<th>N</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Tis</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>I</td>
<td>T1</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIA</td>
<td>T3</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIB</td>
<td>T4a</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIC</td>
<td>T4b</td>
<td>N0</td>
<td>M0</td>
</tr>
<tr>
<td>IIIA</td>
<td>T1-T2</td>
<td>N1/N1c</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T1</td>
<td>N2a</td>
<td>M0</td>
</tr>
<tr>
<td>IIIB</td>
<td>T3-T4a</td>
<td>N1/N1c</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T2-T3</td>
<td>N2a</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T1-T2</td>
<td>N2b</td>
<td>M0</td>
</tr>
<tr>
<td>IIIC</td>
<td>T4a</td>
<td>N2a</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T3-T4a</td>
<td>N2b</td>
<td>M0</td>
</tr>
<tr>
<td></td>
<td>T4b</td>
<td>N1-N2</td>
<td>M0</td>
</tr>
<tr>
<td>IVA</td>
<td>Any T</td>
<td>Any N</td>
<td>M1a</td>
</tr>
<tr>
<td>IVB</td>
<td>Any T</td>
<td>Any N</td>
<td>M1b</td>
</tr>
</tbody>
</table>

### Appendix 3: Breast Cancer TNM Staging System

<table>
<thead>
<tr>
<th>Primary Tumour (T)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TX</td>
<td>Primary tumour cannot be assessed</td>
</tr>
<tr>
<td>T0</td>
<td>No evidence of primary tumour</td>
</tr>
<tr>
<td>Tis</td>
<td>Carcinoma in situ</td>
</tr>
<tr>
<td>T1</td>
<td>Tumour ≤20mm in greatest dimension</td>
</tr>
<tr>
<td>T1mi</td>
<td>Tumour ≤1mm in greatest dimension</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumour &gt;1mm but ≤5mm in greatest dimension</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumour &gt;5mm but ≤10mm in greatest dimension</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>T1c</td>
<td>Tumour &gt;10mm but ≤20mm in greatest dimension</td>
</tr>
<tr>
<td>T2</td>
<td>Tumour &gt;20mm but ≤50mm in greatest dimension</td>
</tr>
<tr>
<td>T3</td>
<td>Tumour &gt;50mm in greatest dimension</td>
</tr>
<tr>
<td>T4</td>
<td>Tumour of any size with extension to the chest wall or skin</td>
</tr>
<tr>
<td>T4a</td>
<td>Extension to chest wall, including pectoralis muscle adherence/invasion</td>
</tr>
<tr>
<td>T4b</td>
<td>Ulceration or ipsilateral satellite nodules and/or oedema of the skin.</td>
</tr>
<tr>
<td>T4c</td>
<td>Both T4a and T4b</td>
</tr>
<tr>
<td>T4d</td>
<td>Inflammatory carcinoma</td>
</tr>
</tbody>
</table>

**Regional lymphnodes (N)**

<table>
<thead>
<tr>
<th>NX</th>
<th>Regional lymphnodes cannot be assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>N0</td>
<td>No regional lymphnode metastasis</td>
</tr>
<tr>
<td>N1</td>
<td>Metastasis to mobile ipsilateral level I/II axillary lymphnodes</td>
</tr>
<tr>
<td>N2</td>
<td>Metastases to ipsilateral level I/II lymphnodes that are clinically fixed</td>
</tr>
<tr>
<td>N2a</td>
<td>Metastases in ipsilateral level I/II axillary lymph nodes fixed to one another or to other structures</td>
</tr>
<tr>
<td>N2b</td>
<td>Metastases only in clinically detected ipsilateral internal mammary nodes and in the absence of clinically evident level I, II axillary lymph node metastases</td>
</tr>
<tr>
<td>N3</td>
<td>Metastases in ipsilateral infraclavicular (level III) lymph nodes, with or without level I, II axillary node involvement, or in clinically detected ipsilateral internal mammary lymph nodes and in the presence of clinically evident level I, II axillary lymph node metastasis; or metastasis in ipsilateral supraclavicular lymph nodes, with or without axillary or internal mammary lymph node involvement</td>
</tr>
<tr>
<td>N3a</td>
<td>Metastasis in ipsilateral infraclavicular lymph nodes</td>
</tr>
<tr>
<td>N3b</td>
<td>Metastasis in ipsilateral internal mammary lymph nodes and axillary lymph nodes</td>
</tr>
<tr>
<td>N3c</td>
<td>Metastasis in ipsilateral supraclavicular lymph nodes</td>
</tr>
</tbody>
</table>

**Distant Metastases (M)**

<table>
<thead>
<tr>
<th>M0</th>
<th>No clinical or radiographic evidence of distant metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>Distant detectable metastases</td>
</tr>
</tbody>
</table>
## Appendix 4: Breast cancer tumour stage as determined by TNM classification groupings

<table>
<thead>
<tr>
<th>Stage 0</th>
<th>Tis, N0, M0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage IA</td>
<td>T1, N0, M0</td>
</tr>
<tr>
<td>Stage IB</td>
<td>T0 or T1, N1mi (micrometastases in 1-3 nodes), M0</td>
</tr>
<tr>
<td>Stage IIA</td>
<td>T0 or T1, N1 (but not N1mi), M0</td>
</tr>
<tr>
<td>Stage IIB</td>
<td>T2, N0, M0</td>
</tr>
<tr>
<td></td>
<td>T2, N1, M0</td>
</tr>
<tr>
<td></td>
<td>T3, N0, M0:</td>
</tr>
<tr>
<td>Stage IIIA</td>
<td>T0 to T2, N2, M0</td>
</tr>
<tr>
<td></td>
<td>T3, N1 or N2, M0</td>
</tr>
<tr>
<td>Stage IIIB</td>
<td>T4, N0 to N2, M0</td>
</tr>
<tr>
<td>Stage IIIC</td>
<td>Any T, N3, M0</td>
</tr>
<tr>
<td>Stage IV</td>
<td>Any T, any N, M1</td>
</tr>
</tbody>
</table>

## Appendix 5: Correspondence between Laura Muirhead and Steve Pringle of Waters (bold) regarding electrical facilities within the operating theatres of St Mary’s Hospital to facilitate development of the Synapt G2-S

**What is the pressure of the compressed gas in the theatres?** The theatre is provided with 4 bar medical air and 7 bar surgical air.

With regards to the mains supply:

1. **Are the wall sockets on separate breakers or are they on a common breaker?** They are grouped in zones

2. **What is the breaking current and time characteristic of the breaker(s) involved?** Traditional breaking capacities of 20A and 32A protective devices. Disconnection times will be in line with BS7671:2008 which are 0.4 seconds socket outlets and 5 seconds for fixed equipment

3. **Is there a RCD/ELCB breaker in line too? If so RCBO’s will be used on certain socket outlets in line with BS 7671:2008**
   a) **is it common or is there one per socket?** There will be one per group of sockets
   b) **what is its rating(s) (5mA/15mA/30mA etc)?** 30mA
   c) **Are there any other devices run on the same breakers (i.e. that might cause a hazard if the MS caused a breaker trip)?** Socket circuits feeding computers are on different breakers than those for fixed equipment or cleaners sockets and vice versa. The NMR equipment and computer sockets will also be backed up by UPS.
Do you know whether the hospital has a forklift that can be used to unload the item Tuesday noon? Unfortunately there is no forklift at St Mary’s. Someone in the Hospital will probably be needing to receive the device, that means as a minimum an acceptance of delivery, but - and I am unaware of local regulations - a technical check may also be required.

Both myself and the lab manager Ken Miller will be available on Tuesday to accept delivery. Neither of us are aware of any technical checks that should be carried out but if they are necessary they should probably be done by a member of our research team who is familiar with MS equipment as no one based in St Mary’s has any knowledge of the equipment.

MediMass may be required to present certain documents we are not aware of, could you give us a contact to ask about this paperwork, or ask about the documents? does the safety office or I don’t know what you have there want to check the machine?

Again, I have asked the lab manager and he is not aware of any documents that are required but I will ask Zoltan and get back to you. I know this is quite unclear yet, but it would be great to know what the device will be used for in which hospital - that is a schedule for the next couple of months would be great. The most likely course of action is that the instrument will be used in St Mary’s until things are set up at the Hammersmith Hospital and then it will be transferred there. This will take at least a few weeks.

We need a 75 cm wide door for the instrument
The route to the 4th floor has doors that are wide enough to accommodate a 75cm wide load, as will the goods lifts and the door to the lab is wide enough.

We need a 10A power outlet. Do you have European power outlet in the OR or only English type?
The power supply in the 4th floor lab and theatres have the normal UK 13amp sockets. Only UK sockets available

Do you have pressurized air in the OR?
There is pressurised air in the 4th floor lab. 2 supplies, 1 for the NMR and 1 for the MS. Photos and descriptions to follow. The fitting is a 13mm copper compression fitting.

Do you have internet in the OR. Wifi or can we use cable if yes. How long cable do we need.
There is no Wifi but the lab has multiple sockets for telephone line/ internet connection close to where the machine will be located. Max length cable required probably 6m and this is a generous estimate to accommodate any instrument/PC setup in the room.

Do we need a transfer letter for the instrument?
Again, unfortunately myself and Ken are really unsure about this. I will ask Zoltan.

Snare Cable
Within the endoscopy department they use the ERBE system for diathermy and APC. The 2 cables, one which connect the power supply to the snare and the other which connects to the patient return electrode come with the instrument and are used for every case. They are only replaced if broken. Having looked at them on the ordering system, the connector for the return electrode is compatible with any instrument therefore we can just continue to use our own setup. However, I’m not sure whether the snare cable is only suitable for use with ERBE. In A&E they have an older endoscopic stack for emergencies and it has what looks like a normal generator but it is also ERBE, although I’m pretty sure the connections are interchangeable. I have emailed the company to ask whether the cable is compatible with another setup and how I can order one if it is compatible.
The price is £65.00. The code is 20192-117 (for reference)

MS Delivery
I have discussed with Julia the practicalities of delivery of the new MS to the 4th floor next Tuesday. She sent a list of questions, most of which I have answered with the help of Ken, however there are a few questions neither of us where sure about.

1. The need for technical check on delivery of instrument. I’m not sure if this is necessary or if it is whether someone from the team should do it as there is no one at St Mary’s who has any idea about the instrument.
2. Do we need a transfer letter with the instrument? MediMass may be required to present certain documents. Both myself and Ken will be at St Mary’s on Tuesday and can deal with the delivery side. Cristina, I think you are away?

Breast Histology
Mike Osborne no longer does breast so has advised I contact Dr Ramakrishnan to see if she will help. Email sent.

Endoscopy Ethics
Ethics amendment and new patient info sheet written. Awaiting input from Zoltan/James to ensure everything is covered before I submit.

MHRA
Phone discussions with MHRA re. handpiece so far unproductive. Have gone down email line instead so we have a record on what is being said. Awaiting response.

Appendix 8: Minutes of iKnife team meetings following installation of first mass spectrometer delivered from Waters into laboratory space at St Mary’s Hospital

iKnife research group meeting 12/04/12 – minutes (distributed by LM).

1. **Handpiece.** Prototype piece looks good. Agreed amendments
   a. 2 extensions from handpiece to be joined together for 1m
   b. More flexible tubing if possible to improve handling
   c. Development of long tip.
2. **Waters instrument** is now ready. LM to send study protocol to Ken Miller for preparation of 10th floor lab space. LM to prepare study design and present at Monday meeting to inform surgeons of study design start date and plan for data collection.
3. **Laparoscopic instrument development.** JK will discuss with Prof Darzi and Karen Kerr regarding access to laparoscopic instruments and development of instrument within department. In meantime we will try to get some kit from the education centre for use at SK.
4. **Endoscopic instrument development.** LM to consider clinical indications e.g. colonic polyps/early rectal ca, Barretts, IBD. Once we have a plan we will speak to Dr Orchard (gastroenterologist). Need to check ethics regarding endoscopy and amend accordingly. Christian/Julia to have a look at endoscopic regarding possible amendments for use once Laura has access to scope.
5. **Access to theatre** for Julia, Christian, Cristina. LM will discuss with Ann McKenna and pass on requirements for access to team.
6. **Biobank tissue** for ex vivo analysis. LM to find out what tissue is available for use, how it is stored etc.

7. **Publications.** Most of the text written. ZT to send draft to Prof Nicholson to look at over the weekend. He will then pass on to rest of team. The top copy only will be sent. If amendments are made, please ensure the document is saved with information on who has edited it and when and send on. Please copy in rest of team when forwarding document.

8. **NIHR translational research programme grant.** Deadline 25th May. This should be very clinically orientated, focusing on 1 clinical problem e.g. management of early colorectal cancer/lymphnode involvement/local spread. This could be based around both the colorectal NMR data and the iKnife data. JK will send outline.

---

**Iknife research group meeting 04.05.2012 – minutes (Distributed by LM)**

Instrument installed in lab space. Size and separate PC make it unsuitable for theatre use.

1. Sensitivity vs resolution. Resolution seems good. 20 000. Objective to classify samples. If using multivariate methods for tissue classification then sensitivity is crucial. To ID biomarkers resolution is crucial. With current objective should focus on sensitivity. Is it worth increasing resolution of we lose sensitivity? We could do some studies getting some model systems and get systematic data regarding resolution practically and looking at what the improved resolution actually gives us. Higher res gives more peaks but whether they correspond to something important or not is what we need to find out. We should design a study to see whether improved resolution gives us more information.  

2. LM to go with RM to look at sectioning of colonic tissue. Mass spec study of heterogeneity of normal colonic mucosa. Try and get as much tissue as we can. We can run 10 samples/day on the machine.

3. LM to get theatre lists.

4. Pressure required for venturi roughly 5 bars.

5. CG to do some electrospray/nanospray early next week. Julia dry blot spot extracts. Source parts should arrive on Thursday. CS to look for 0.53 tubing (stainless steel)

6. In one week the instrument should be up and running with REIMS spectra with acceptable quality. Then there are 2 factors  
   1. Optimisation of instrument ie changing front end.  
   2. Obtaining data

7. Laura to source endoscope snare.