

Delineating the difficulties in diagnosing
oesophageal adenocarcinoma destined to
arise from Barrett's oesophagus and
exploring the role of vibrational
spectroscopy from biofluid and tissue
analysis

School of Pharmacy and Biomedical Sciences

MD

Ishaan Maitra

2020



Delineating the difficulties in diagnosing
oesophageal adenocarcinoma destined to
arise from Barrett's oesophagus and
exploring the role of vibrational
spectroscopy from biofluid and tissue
analysis

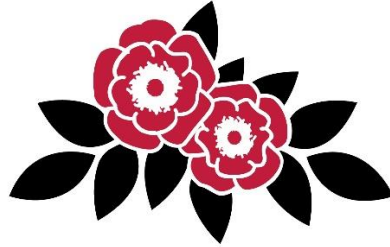
Mr Ishaan Maitra

MD (Res)

Funded by the Rosemere Foundation

STUDENT DECLARATION FORM

uclan



Concurrent registration for two or more academic awards Either ***I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution**

or *I declare that while registered for the research degree, I was with the University's specific permission, a *registered candidate/*enrolled student for the following award:

Material submitted for another award


Either ***I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work**

or *I declare that the following material contained in the thesis formed part of a submission for the award of

(state award and awarding body and list the material below):

Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

Signature of Candidate  Ishaan Maitra

Type of Award MD Res

School School of Pharmacy and Biomedical sciences

Abstract

Histological analysis is the current gold standard in diagnosing Barrett's oesophagus and oesophageal cancer. In this thesis, we look at the feasibility of vibrational spectroscopy as a safe, diagnostic alternative. It can limit the need for endoscopy thus reduce the financial burden on the National Health Service (NHS) as well as reduce patient anxiety by minimising invasive testing.

Studies aimed at targeting both biofluid and tissue diagnosis of Barrett's oesophagus and stages to oesophageal adenocarcinoma (OAC) have been performed. Spectrochemical techniques including Attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR) and Raman spectroscopy were coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA). We have produced promising results suggesting that this is a powerful alternative in the detection of all oesophageal stages of disease to OAC.

For biofluids, ATR-FTIR spectroscopy on plasma using the GA-QDA model demonstrated excellent accuracy, sensitivity and specificity for all classifications of oesophageal disease to OAC. The accuracy, sensitivity and specificity were shown to be 100% for all classifications of oesophageal disease to OAC where the PCA-QDA model was utilised on urine samples. Raman spectroscopy on saliva and urine using the GA-QDA model obtained an accuracy of 100% for all classifications to OAC.

For tissue, ATR-FTIR spectroscopy on *ex vivo* samples using the SPA-QDA model proved superior compared to PCA-QDA and GA-QDA models, demonstrating accuracy >90% for all classifications of oesophageal disease to OAC. Raman spectroscopy on *ex vivo* samples established >71.4% for all figures of merit in all classes of oesophageal disease to OAC using the PCA-QDA model. The figures of merit for distinguishing LGD and OAC were specifically excellent (100%).

An additional unique study in 3 index patients using Raman tissue mapping has identified subtle spectral discriminant markers such as β -glucose, lipids, collagen and protein peaks found in normal squamous epithelial tissue and specialised intestinal metaplasia which are elevated in OAC tissue. These markers may suggest the probability of benign oesophageal pathology transforming into OAC in the future. These patients may warrant further intensive surveillance.

In conclusion, this thesis supports previous literature that ATR-FTIR and Raman spectroscopy may be utilised in the diagnostic as well as surveillance processes for Barrett's oesophagus and its progression to OAC. The novel finding of biofluids being able to classify disease progression to OAC is pertinent. This would require further multicentre study for validation.

Contents

Student Declaration	II
Abstract	III
Acknowledgements	VI
List of Figures	VII
List of Tables	XI
List of Abbreviations	XIII
1.Introduction	2
1.1 Barrett’s oesophagus	3
1.1.1 Historical perspective and Definition	4
1.1.2 Endoscopic evaluation	5
1.1.3 Risk factors and Epidemiology	7
1.1.4 Natural History and Progression to OAC	8
1.1.5 Histopathology	10
1.1.6 Molecular and Genetic advances	15
1.1.7 Field Effect	17
1.1.8 Recognised Biomarkers	18
1.1.9 Screening	21
1.1.10 Surveillance	23
1.1.11 Management	26
1.1.12 Summary	28
1.2 Advanced imaging technology in Barrett’s oesophagus	29
1.2.1 High resolution endoscopy	29
1.2.2 Chromoendoscopy	29
1.2.3 Autofluorescence	30
1.2.4 Narrow band imaging	30
1.2.5 Optical coherence tomography	30
1.2.6 Confocal fluorescence microendoscopy	31
1.3 Vibrational spectroscopy	32
1.3.1 Principles	32
1.3.2 Attenuated total reflection / Fourier Transform infrared Spectroscopy	33

1.3.3 Raman Spectroscopy	35
1.3.4 Properties of biofluids and tissue in spectroscopy	36
1.3.4.1 Biofluid analysis	36
1.3.4.2 Human tissue analysis	37
1.4 The role of biospectroscopy in cancer diagnostics	38
1.5 Conclusion	40
1.6 Aims of Thesis	41
1.6.1 Objectives	41
2. Study demographics	42
3. Attenuated total reflection Fourier-transform infrared spectral discrimination in oesophageal transformation to adenocarcinoma from human body fluids	44
4. Raman spectral discrimination in human liquid biopsies of oesophageal transformation to adenocarcinoma	69
5. Attenuated total reflection Fourier-transform infrared spectral discrimination in oesophageal transformation to adenocarcinoma from human tissue	95
6. Discrimination of oesophageal transformation to adenocarcinoma on human tissue samples using Raman microscopy	109
7. Establishing spectrochemical changes in the natural history of oesophageal adenocarcinoma from tissue Raman mapping analysis	123
8. Discussion	148
8.1 Further work	152
9. References	153
10. Appendix	171
10.1 Supplementary tables	171
10.2 Patient Consent form (Biofluids and Tissue)	181
10.3 Cover letter approval for Ethics (HRA)	182
10.4 Cover letter approval for Ethics (UCLan)	186
10.5 Publications	187
10.6 Presentations	187

Acknowledgements

I have sincerely enjoyed my journey completing this MD. I am indebted to both Professor Francis L. Martin and Mr Ravindra S. Date for their ongoing guidance and reassurance. I am very grateful to both of you for not only giving this opportunity to me but keeping me on track through the years.

I would like to thank Kate Ashton at Lancashire Teaching Hospitals NHS Foundation Trust. You were always on hand with advice at all times of the day. A mention also goes to Danielle Bury, who helped with the tissue analysis of the research. Furthermore, she helped me to understand basic histopathology again!

To Camilo and Kassio - you have both helped me tremendously. I am forever grateful for your patience with me whilst I undertook the project.

Maria - a big thank you for answering the multitude of questions I asked of you initially in this process, and for your help in using the Raman technology in particular. Panos – you also kept me going through the whole process, a big thanks for all your moral support.

My parents and wife Mitali have always supported me through this MD. They have always been at my side.

Finally, I would like to sincerely thank all the patients who were recruited onto this study as well as the Rosemere Cancer Foundation. Without your generosity and understanding in donating your tissue and biofluid samples this MD would never have been possible.

List of Figures

Figure 1: Illustration showing ‘C’ and ‘M’ measurements

Figure 2: Endoscopic visualisation of A) Barrett’s oesophagus segment B) Barrett’s oesophagus with LGD and HGD C) OAC

Figure 3: Photomicrograph of Intestinal Metaplasia (H&E \times 100 objective)

Figure 4: Photomicrograph of Low-Grade Dysplasia (H&E \times 100 objective)

Figure 5: Photomicrograph of High-grade dysplasia (H&E \times 100 objective)

Figure 6: Photomicrograph of Oesophageal adenocarcinoma (H&E \times 100 objective)

Figure 7: Surveillance flow chart for non-dysplastic Barrett’s oesophagus

Figure 8: Surveillance and management for dysplasia (LGD/HGD)

Figure 9: Electromagnetic radiation spectrum

Figure 10: FTIR spectrometer diagram

Figure 11: Rayleigh, stokes and Anti-stokes scattering

Figure 12: Examples of biofluid slides for ATR-FTIR spectroscopy

Figure 13: Bruker TENSOR 27 FTIR spectrometer

Figure 14: Comparison of normal/inflammatory/Barrett’s oesophagus/LGD/HGD/OAC oesophageal stages using plasma samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett’s oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 15 selected wavenumbers by GA-QDA. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.

Figure 15: Comparison of normal/inflammatory/Barrett’s oesophagus/LGD/HGD/OAC oesophageal stages using saliva samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett’s oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) Singular value decomposition (SVD) vs. Number of principal component (PC) obtained by PCA-QDA, where Power represents the eigenvalue. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for PCA-QDA. \circ - illustrates a misclassification.

Figure 16: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using serum samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 13 selected wavenumbers by GA-QDA. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 17: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using urine samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) Singular value decomposition (SVD) vs. Number of principal component (PC) obtained by PCA-QDA, where Power represents the eigenvalue. (D) Predicted class plot vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for PCA-QDA. ○ - illustrates a misclassification.

Figure 18: Example of biofluid slide for Raman spectroscopy

Figure 19: Renishaw 100 System

Figure 20: Graphical abstract demonstrating how the four biofluids (plasma, serum, urine and saliva) were measured through Raman spectroscopy in order to detect oesophageal transformation stages to adenocarcinoma

Figure 21: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using plasma samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 16 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 22: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using saliva samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 16 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 23: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using serum samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 30 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 24: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using urine samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 29 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 25: Graphical abstract demonstrating how oesophageal tissue is processed through ATR-FTIR spectroscopy in order to detect oesophageal transformation stages to adenocarcinoma

Figure 26: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using tissue samples for ATR spectroscopy. The panel shows: (A) Average raw spectra in the ATR region of 1800 cm^{-1} to 900 cm^{-1} and (B) Average preprocessed ATR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 16 selected variables used by GA-QDA model. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 27: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using tissue samples for Raman spectroscopy. The panel shows: (A) Average raw spectra in the Raman region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average preprocessed Raman spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 16 selected variables by the GA-QDA model. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. ○ - illustrates a misclassification.

Figure 28: Ulcerative lesion on background of Barrett's oesophagus (35 cm from incisors, 1 cm length).

Figure 29: Graphical abstract demonstrating how oesophageal tissue is processed through Raman spectroscopy and mapped in order to detect oesophageal transformation stages to adenocarcinoma.

Figure 30: (a) Average raw Raman spectra for OAC and normal tissue (case 1); (b) average pre-processed (Savitzky-Golay smoothing [21 points window, 2nd order polynomial fitting] and AWLS baseline correction) Raman spectra for OAC and normal tissue (case 1); (c) PC scores plot for OAC and normal tissue; (d) PCA-LDA discriminant function (DF) plot for OAC and normal tissue (case 1), where o = training set and * = validation set.

Figure 31: (a) Raw and (b) PCA-recovered images for normal tissue; (c) raw and (d) PCA-recovered images for OAC tissue; (e) difference-between-mean (DBM) spectrum and PC loadings between normal tissue vs. OAC (case 1). Colour bar: mean relative intensity.

Figure 32: (a) Average raw Raman spectra for Barrett's oesophagus and OAC tissue (case 2); (b) average pre-processed (Savitzky-Golay smoothing [21 points window, 2nd order polynomial fitting] and AWLS baseline correction) Raman spectra for Barrett's oesophagus and OAC tissue (case 2); (c) PC scores plot for Barrett's oesophagus tissue and OAC tissue; (d) PCA-LDA discriminant function (DF) plot for Barrett's oesophagus and OAC tissue (case 2), where o = training set and * = validation set.

Figure 33: (a) Raw and (b) PCA-recovered images for Barrett's oesophagus tissue; (c) raw and (d) PCA-recovered images for OAC tissue; (e) difference-between-mean (DBM) spectrum and PC loadings between Barrett's oesophagus vs. OAC tissue (case 2). Colour bar: mean relative intensity.

Figure 34: (a) Average raw Raman spectra for Barrett's oesophagus and OAC tissue (case 3); (b) average pre-processed (Savitzky-Golay smoothing [21 points window, 2nd order polynomial fitting] and AWLS baseline correction) Raman spectra for Barrett's oesophagus and OAC tissue (case 3); (c) PC scores plot for Barrett's oesophagus and OAC tissue (case 3); (d) PCA-LDA discriminant function (DF) plot for Barrett's oesophagus and OAC tissue (case 3), where o = training set and * = validation set.

Figure 35: (a) Raw and (b) PCA-recovered images for Barrett's oesophagus tissue; (c) raw and (d) PCA-recovered images for OAC tissue; (e) difference-between-mean (DBM) spectrum and PC loadings between Barrett's oesophagus vs. OAC tissue (case 3). Colour bar: mean relative intensity.

List of Tables

Table 1: Worldwide professional societies' definitions of Barrett's oesophagus

Table 2: Vienna classification of epithelial neoplasia of the digestive tract

Table 3: Summary of molecular biomarkers

Table 4: Patient demographics

Table 5: Number of training, validation and prediction specimens (or spectra) in each category of biofluids

Table 6: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA and SPA-QDA) results for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using plasma samples

Table 7: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using saliva samples

Table 8: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using serum samples

Table 9: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using urine samples

Table 10: Number of training, validation and prediction samples in each category of biofluids

Table 11: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using plasma samples.

Table 12: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using saliva samples.

Table 13: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using serum samples.

Table 14: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using urine samples.

Table 15: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using tissue samples.

Table 16: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for the multivariate classification models (PCA-QDA, SPA-QDA and GA-QDA) to distinguish normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using tissue samples.

Table 17: Quality parameters (accuracy, sensitivity and specificity) for distinguishing normal *vs.* OAC (case 1) tissue using PCA-LDA.

Table 18: Quality parameters (accuracy, sensitivity and specificity) for distinguishing Barrett's oesophagus vs. OAC (case 2) tissue using PCA-LDA.

Table 19: Quality parameters (accuracy, sensitivity and specificity) for distinguishing Barrett's oesophagus vs. OAC (case 3) tissue using PCA-LDA.

List of abbreviations

ACG	American College of Gastroenterology
AGA	American Gastroenterological Association
ALS	Asymmetric least squares
ATP	Adenosine Triphosphate
ATR	Attenuated total reflection
AWLS	Automatic weighted least squares
BMI	Body Mass Index
BO	Barrett's Oesophagus
BOSS	Barrett's Oesophagus Surveillance Study
BSG	British Society of Gastroenterology
CA-125	Cancer Antigen 125
cAMP	Cyclic adenosine monophosphate
CCD	Charge-coupled device
CLO	Campylobacter-like organism
COPD	Chronic obstructive pulmonary disease
CREB	cAMP Response Element-Binding protein
CRTC1	CREB-regulated transcription coactivator 1
CT	Computerised Tomography
DBM	Difference Between Mean
DI	Diaphragmatic Indentation
DNA	Deoxyribonucleic acid
EET	Endoscopic Eradication Therapy
EMR	Endoscopic Mucosal Resection
FOM	Figures of Merit
FOXF1	Forkhead box protein F1
FOXP1	Forkhead box protein P1
FTIR	Fourier-transform infrared spectroscopy
GA	Genetic algorithm

GA-QDA	Genetic algorithm quadratic discriminant analysis
GFR	Glomerular Filtration Rate
GI	Gastrointestinal
GOJ	Gastro-oesophageal junction
GORD	Gastro-oesophageal reflux disease
H&E	Haematoxylin and Eosin
HCA	Hierarchical cluster analysis
HCC	Hepatocellular carcinoma
HGD	High-grade dysplasia
HRA	Health Research Authority
IHC	Immunohistochemistry
IL	Interleukin
IM	Intestinal Metaplasia
IR	Infrared
LDA	Linear Discriminant Analysis
LGD	Low-grade dysplasia
LOH	Loss of Heterozygosity
McM2	DNA replication licensing factor MCM2
MHC	Major Histocompatibility Complex
NADH	Nicotinamide Adenine Dinucleotide
NBI	Narrow Band Imaging
NFκB	Nuclear Factor kappa-light-chain-enhancer of activated B cells
NHS	National Health Service
NICE	National Institute of Clinical Excellence
OAC	Oesophageal adenocarcinoma
OCT	Optical coherence tomography
OGD	Oesophagogastroduodenoscopy
OR	Odds ratio
PAS	Periodic acid-Schiff

PCA	Principal component analysis
PCA-LDA	Principal component analysis- linear discriminant analysis
PCA-QDA	Principal component analysis- quadratic discriminant analysis
PDT	Photodynamic Therapy
QALY	Quality Adjusted Life Year
QDA	Quadratic discriminant analysis
REC	Research Ethics Committee
RFA	Radiofrequency Ablation
RNA	Ribonucleic acid
ROC	Receiver Operating Characteristic
RS	Raman Spectroscopy
SCC	Squamous cell carcinoma
SFED	Société Française d'Endoscopie Digestive
SIM	Specialised Intestinal Metaplasia
SPA-QDA	Successive projections algorithm quadratic discriminant analysis.
SURF	Surveillance with Radio-Frequency Ablation
SVM	Support Vector Machine
TFF3	Trefoil Factor 3
UK	United Kingdom
USA	United States of America

Chapter

1

1. Introduction

The diagnosis of oesophageal adenocarcinoma (OAC) has serious consequences given its late diagnosis and poor prognosis. Although survival rates are poor, a better survival rate has been observed amongst patients with adenocarcinoma arising from Barrett's mucosa compared with oesophageal malignancies without metaplasia (Johansson *et al.*, 1996; Thomas *et al.*, 1997).

The genetic heterogeneity and complexity of advanced cancers such as OAC strongly support the rationale for early interruption of the carcinogenic process (Meyskens *et al.*, 2016). The emphasis of cancer prevention management should be on individuals at high risk and on primary localised disease in which screening and surveillance should play a paramount role.

Tissue biopsy and histopathological confirmation remains the gold standard in cancer diagnostics. Tissue biopsies in OAC involve invasive endoscopy techniques, which can be painful and uncomfortable for patients. Furthermore, intra and inter-observer error exists in the reporting of specimens with dysplasia. This may require specialist histopathological input which can subsequently delay a cancer diagnosis and exacerbate patient anxiety. Novel innovative diagnostic tools are necessary to help histopathologists in ensuring accurate and reproducible diagnostics whilst decreasing turnaround times.

The aim of this thesis is to assess the feasibility of vibrational spectroscopy in Barrett's oesophagus and in all classifications to OAC, as a safe, diagnostic alternative.

1.1 Barrett's oesophagus

The rising incidence of oesophageal cancer over the past three decades coincides with a change in the histological type and primary tumour location (Pohl *et al.*, 2005). Oesophageal adenocarcinoma (OAC) has now replaced squamous cell carcinoma (SCC) as the most common type of oesophageal malignancy in the western world (Blot *et al.*, 1999). OAC is aggressive and usually presents late with a poor prognosis with an overall 5-year survival below 25% (Zagari *et al.*, 2008). Even in spite of technological enhancements related to preventative strategies and more effective combination therapies, the overall incidence of oesophageal adenocarcinoma has risen (Layke *et al.*, 2006).

There is a proven association between oesophageal adenocarcinoma and Barrett's oesophagus, a condition that appears to arise in response to chronic inflammation from gastro-oesophageal reflux disease (GORD) (Wong *et al.*, 2005). Reflux induces metaplasia, which in turn leads to high grade dysplasia (HGD) and invasive OAC. Barrett's oesophagus is the only known precursor to OAC to date, and has a small prevalence of up to 2% in the European population (Ronkainen *et al.*, 2005).

General surveillance through oesophagogastroduodenoscopy (OGD) of all individuals with Barrett's oesophagus is not cost-effective as the annual incidence of oesophageal adenocarcinoma developing in Barrett's oesophagus is only 0.33% (Desai *et al.*, 2012; Gordon *et al.*, 2014). This highlights the need to adapt surveillance programs to include individuals with Barrett's oesophagus at high absolute risk of tumour progression.

There still remains a lack of consensus regarding the natural history and definition in Barrett's oesophagus (Spechler *et al.*, 2011). Furthermore, there is a lack of reliable predictive biomarkers that might enable us to risk-stratify Barrett's oesophagus patients and identify those who would benefit the most from endoscopic management (Moyes *et al.*, 2011; Bhardwaj *et al.*, 2012). Finally, prospective studies and present guidelines have not established a clear survival and cost-effectiveness benefit for screening and surveillance where robust risk stratification can be utilised by clinicians for patients to best use NHS resources (Spechler *et al.*, 2011; Amadi *et al.*, 2017).

1.1.1 Historical perspective and Definition

Australian born surgeon Norman Rupert Barrett first coined the term ‘Barrett’s oesophagus’ in patients with ulcerative changes in a tubular organ suggestive of an oesophagus. He clarified that the distal ulcerated portion was lined by columnar epithelium (Gindea *et al.*, 2014). Boshier *et al.* (1951) elaborated on Mr Barrett’s definition describing intestinal type goblet cells in the columnar-lined oesophagus.

The first suggested case of malignancy related to Barrett’s oesophagus was reported by Morson and Belcher in 1952. The authors reported the case of a patient who developed an adenocarcinoma in oesophageal mucosa that presented ‘atrophic changes with a tendency towards intestinal type tissue containing many goblet cells’ (Morson *et al.*, 1952).

The general consensus by the end of the 1970’s was that intestinal metaplasia was widely regarded as both the most common type of Barrett’s epithelium and that intestinal metaplasia was associated with cancer development.

The definition of Barrett’s oesophagus is contentious and worldwide professional societies have previously outlined different definitions related to biopsy evidence of intestinal metaplasia (Table 1). Without comprehensive population-based studies it is difficult to define the true incidence of the disease. The UK definition has now incorporated a histological confirmation metaplastic columnar epithelium (Fitzgerald *et al.*, 2014). Specialised intestinal metaplasia characterised by the presence of goblet cells is associated with the risk of progression to low-grade dysplasia (LGD), high-grade dysplasia (HGD) and adenocarcinoma (OAC).

Table 1: Worldwide professional societies’ definitions of Barrett’s oesophagus

	ACG (USA) (Spechler <i>et al.</i> , 2011)	BSG (England) (Playford <i>et al.</i> , 2006)	AGA (USA) (Wang <i>et al.</i> , 2008)	SFED (France) (Boyer <i>et al.</i> , 2007)
Intestinal metaplasia (biopsies)	Yes	No	Yes	Yes
Endoscopic documentation of columnar lined mucosa	Yes	Yes	Yes	Yes

ACG: American College of Gastroenterology; BSG: British Society of Gastroenterology

AGA: American Gastroenterological Association; SFED: Société Française d'Endoscopie Digestive

1.1.2 Endoscopic evaluation

The significant anatomical landmark in Barrett's oesophagus identification is the gastro-oesophageal junction (GOJ). This is typically identified as the proximal extent of the upper gastric folds. The squamocolumnar junction (Z line) is the point at which squamous mucosa of the oesophagus meets the columnar mucosa of the stomach. In the absence of SIM, the Z line and GOJ coincide. The diaphragmatic indentation (DI) can be used to identify the presence of a hiatus hernia where the GOJ lies above the DI.

Barrett's oesophagus was initially characterised into short and long segment disease (< or > 3 cm) based on endoscopic findings (Sharma *et al.*, 1998). Since its presentation in 2006, the Prague classification has been regarded as the standard for measuring the length of Barrett's oesophagus (Sharma *et al.*, 2006 (A)). The Prague 'C' and 'M' classification accounts for the circumferential (C) and maximum (M) tongue extent of endoscopically visualised Barrett's oesophagus above the GOJ.

Figure 1: Illustration showing 'C' and 'M' measurements

Adapted from <https://www.endoscopy-campus.com/en/klassifikationen/prag-klassifikationen-des-barrett-oesophagus/> (accessed 16th October 2019)

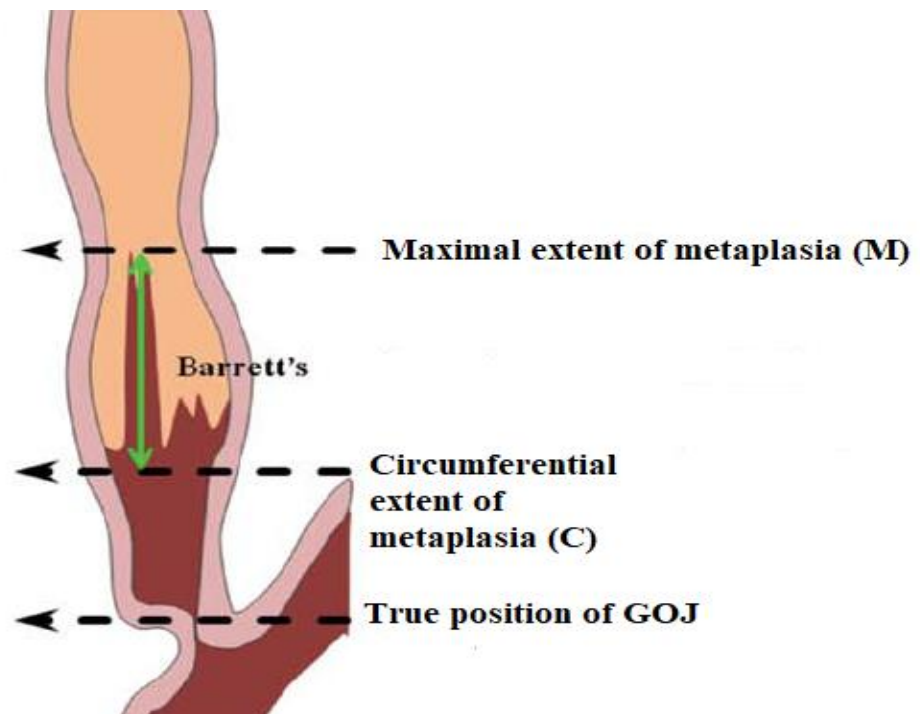
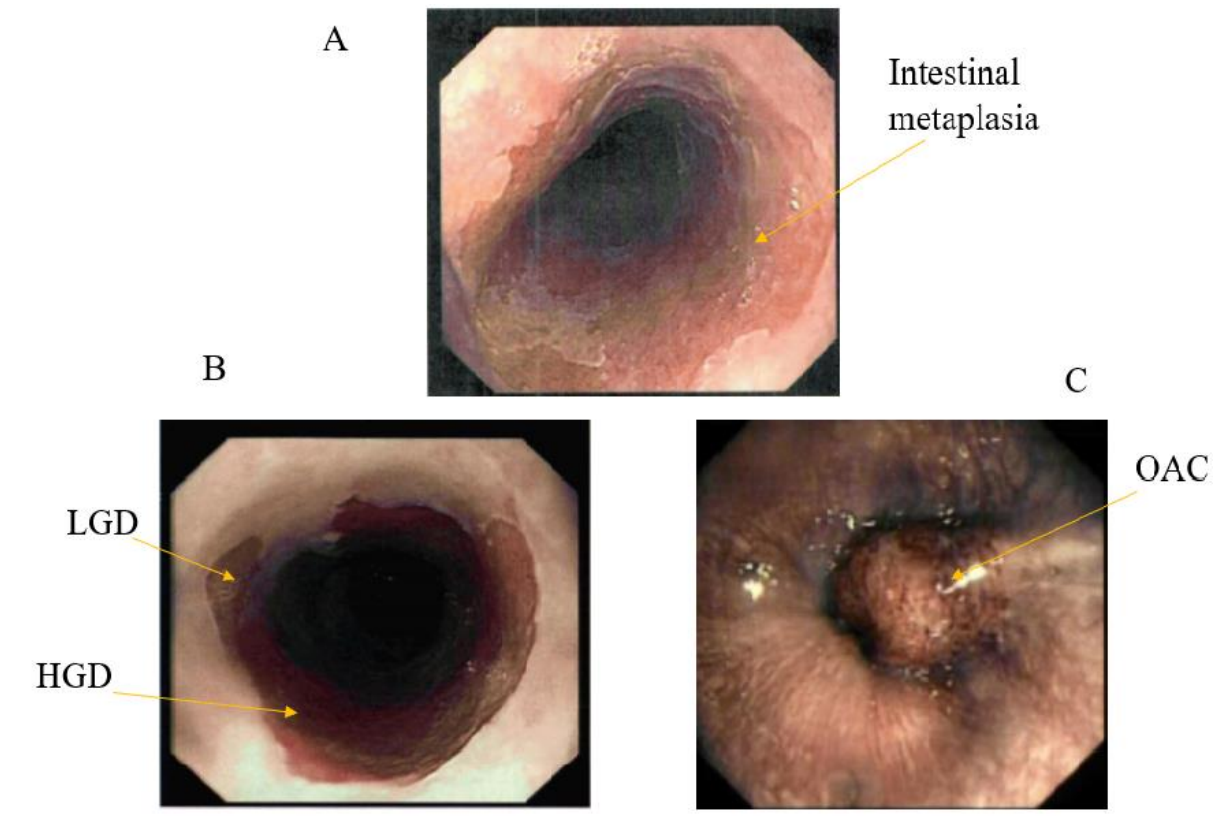


Figure 2: Endoscopic visualisation of A) Barrett's oesophagus segment B) Barrett's oesophagus with LGD and HGD C) OAC



The sensitivity of endoscopic examination for the diagnosis of Barrett's oesophagus increases as the length of the Barrett's oesophagus segment increases (Eloubeidi *et al.*, 1999). The overall prevalence of short-segment Barrett's oesophagus is greater than long-segment Barrett's oesophagus in studies where biopsies are systematically taken using established protocols (Hirota *et al.*, 1999). The overall reliability of endoscopic examination and biopsy is approximately 80% (Hirota *et al.*, 1999). Given the inter-observer variability of diagnosing short segments, especially those < 1 cm, and the exclusion of this length from some studies, it is uncertain whether segments < 1 cm are associated with an increased risk of OAC (Pohl *et al.*, 2016). There may be a greater risk of progression to dysplasia and adenocarcinoma with long-segment specialised intestinal metaplasia, perhaps related to more severe reflux (Pohl *et al.*, 2016).

1.1.3 Risk factors and Epidemiology

Males in their 6th decade of life are twice as likely to have a diagnosis of Barrett's oesophagus than females (Yachimski *et al.*, 2010). White subjects have a 4–6 times higher incidence of the disease compared to black subjects (Corley *et al.*, 2009). Other risk factors such as chronic GORD and smoking increase the risk of Barrett's oesophagus in multiple studies (Corley *et al.*, 2007; Cook *et al.*, 2012). Being infected with *Helicobacter pylori* is inversely associated with Barrett's oesophagus prevalence (Provenzale *et al.*, 1999). Centripetal obesity has been linked to Barrett's oesophagus and progression to malignancy by favouring a pro-inflammatory state by enhanced quantities of interleukin-6 (IL-6) and interleukin-8 (IL-8) (Mudyanadzo, 2018). Alcohol consumption does not appear to be a strong risk factor.

The worldwide increase in GORD is accompanied by a rapid increase in the prevalence of Barrett's oesophagus (Corley *et al.*, 2007). The exact prevalence in different populations is difficult to assess as the condition is asymptomatic and a diagnosis is made only when an endoscopy is performed (Kuipers *et al.*, 2018). The latter is usually performed for persistent GORD symptoms. The prevalence of Barrett's oesophagus in the unselected general population is between 1%-2% in European studies (Zagari *et al.*, 2008) and approximately 5 to 6% in United States studies (Hayeck *et al.*, 2010).

Several population-based studies have provided further insight into the prevalence of Barrett's oesophagus. The Swedish *Kalixanda* study found that up to 10.3% of patients with GORD symptoms ($n = 1000$) had an endoscopic Barrett's oesophagus segment (Ronkainen *et al.*, 2005). In an Italian study of 1033 patients, histology confirmed the presence of specialised intestinal metaplasia in 1.6% of patients (Zagari *et al.*, 2008).

Further UK detailed studies on age- and sex-related distribution of Barrett's oesophagus have observed that the prevalence of the condition increased with 1.4% for each additional year of age between the age of 20 and 59 year in males. It showed a similar pattern with a 20-year delay in the female population (van Blankenstein *et al.*, 2005).

In a meta-analysis of 23 studies from Asia, the prevalence of endoscopically diagnosed Barrett's oesophagus was 7.8% with histologically confirmed Barrett's oesophagus in 1.6% (Shiota *et al.*, 2015). The prevalence of LGD, HGD and adenocarcinoma in cases with histologically proven Barrett's oesophagus was 6.9%, 3.0% and 2.0%, respectively (Shiota *et al.*, 2015). This compares with studies from the United States of America with rates of up to 9.8%; 5.3% and 5.7%, respectively (Singh *et al.*, 2014).

1.1.4 Natural History and Progression to OAC

It has been shown that specialised intestinal metaplasia arising in Barrett's oesophagus is a risk factor for progression to adenocarcinoma. A meta-analysis from 2012 including 57 studies ($n = 11434$ patients) reported an annual incidence of oesophageal adenocarcinoma developing in Barrett's oesophagus of 0.33% (95% CI 0.28–0.38%) (Desai *et al.*, 2012). Patients with Barrett's oesophagus are 10 times more likely to die from other causes than oesophageal cancer such as cardiorespiratory diseases (Desai *et al.*, 2012). Furthermore, it appears that men progress to adenocarcinoma at twice the rate of women, patients with short segment Barrett's oesophagus are least likely to progress to adenocarcinoma and those with dysplasia at index biopsy are the most likely to progress to adenocarcinoma (Desai *et al.*, 2012).

Progression rates to dysplasia and adenocarcinoma were initially established from a Dutch cohort study in Rotterdam. One hundred and sixty-six patients were recruited from 1973 to 1983 and endoscopic surveillance was started from 2001 (den Hoed *et al.*, 2011). Thirteen patients (M/F 10/3; 7.8%) developed HGD or OAC during follow-up. These were all symptomatic cases of HGD/OAC as the patients were not under endoscopic surveillance and were only re-investigated for symptoms of GORD.

Subsequent meta-analyses and systematic reviews from 2007 to 2017 report progression rates to adenocarcinoma ranging from 2.2 and 6.3 per 1000 patient-years when focusing on all patients with Barrett's oesophagus (Thomas *et al.*, 2007; Qiao *et al.*, 2015; Qumseya *et al.*, 2017). Singh *et al.* (2014) conducted a systematic review of 24 studies ($n = 2694$). This concluded an annual progression rate of Barrett's oesophagus to adenocarcinoma of 5.4 (3.2–7.6) per 1000 patient-years, and progression to high-grade dysplasia or adenocarcinoma combined of 17.3 (9.9–24.7) per 1000 patient-years.

The majority of these studies were dependent upon a single pathologist's interpretation of the histology. Since the diagnosis of dysplasia is investigator dependent, more rigorous criteria ask for confirmation of dysplasia by a second pathologist. This often leads to down-staging of a proportion of patients which may affect disease progression rates and epidemiological studies (Kuipers *et al.*, 2018). Kestens *et al.* (2016) re-assessed LGD samples in 231 Barrett's oesophagus patients. LGD was confirmed in 70%; the remainder was mostly downgraded to no dysplasia or indefinite for dysplasia.

A recent systematic review and meta-analysis conducted in 2017 confirmed that the risk of progression to HGD or OAC in Barrett's oesophagus patients was primarily determined by the presence or absence of LGD (OR 4.2 (2.1–8.5)) (Krishnamoorthi *et al.*, 2018).

Patients with GORD can develop Barrett's oesophagus and subsequently go on to develop LGD, HGD and OAC. At an early stage, these conditions can be treated by ablative and minimally invasive techniques with limited risk. However, at an advanced stage, OAC requires invasive treatment with considerable burden, financial cost, and mortality (Kuipers *et al.*, 2018). Early detection and prevention are the key strategies in managing OAC. The argument as to which Barrett's oesophagus patients are most likely to benefit from surveillance and management hinges on the high prevalence of Barrett's oesophagus and the low cancer incidence amongst unselected Barrett's oesophagus cases (Kuipers *et al.*, 2018).

With the overall risk of OAC in Barrett's oesophagus being low, patients are often middle to older ages with obesity and metabolic syndrome. These patients are more susceptible to cardiac or pulmonary events. This is often not conveyed to patients and many patients fear developing malignancy which is actually disproportionate to their actual risk. This further highlights the need to adapt surveillance programs to include individuals with Barrett's oesophagus at high absolute risk of tumour progression.

1.1.5 Histopathology

For a non-suspicious Barrett's oesophagus segment undergoing routine surveillance, mapping biopsies should be taken at 2 cm intervals from each quadrant as well as separate biopsies from the anatomic cardia. For segments suggestive of dysplasia, biopsies should be taken at 1 cm intervals. This so-called Seattle protocol (Sampliner, 1998) increases the yield of both LGD and HGD by 17% and 3%, respectively, compared with random biopsies (Abela *et al.*, 2008). A minimum of eight biopsies are required to provide an acceptable degree of histological confirmation of Barrett's oesophagus (Harrison *et al.*, 2007). In patients with Barrett's oesophagus < 3 cm and no intestinal metaplasia or dysplasia, a repeat endoscopic assessment with quadrantic biopsies is recommended to establish the diagnosis (Fitzgerald *et al.*, 2014).

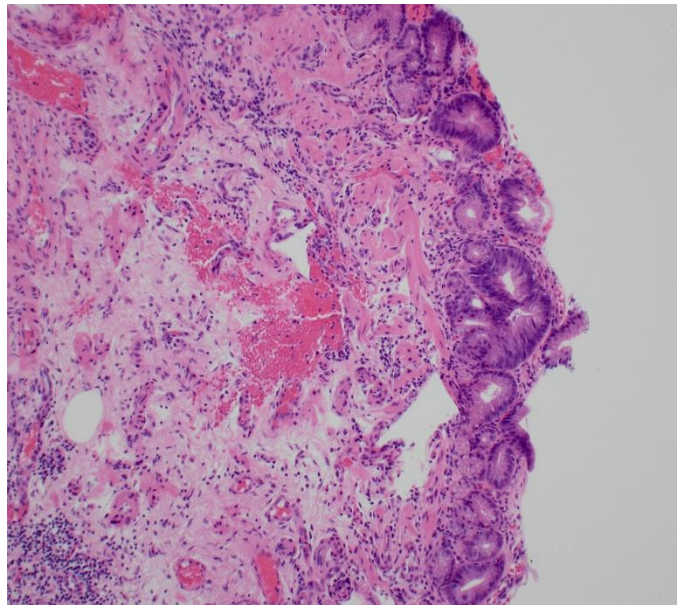
Histopathological definitions of Barrett's oesophagus vary worldwide and are complex (Takubo *et al.*, 2009). The UK requires endoscopic with histological confirmation of metaplastic columnar epithelium \geq 1cm the GOJ (Fitzgerald *et al.*, 2014). The requirement of SIM has been supported by studies that have claimed intestinal metaplasia to be a prerequisite for the development of adenocarcinoma (Ruol *et al.*, 2000).

Goblet cells may be identified on routine histological stains (Haematoxylin and Eosin, H&E), although many organisations routinely employ PAS-Alcian Blue stains to highlight acidic mucin in goblet cells. This stain is helpful in differentiating true SIM from so-called columnar blue cells or pseudo-goblet cells, which are gastric foveolar cells that have been affected by GORD (Bhardwaj *et al.*, 2012).

Intestinal metaplasia of the gastric cardia does not compare to Barrett's oesophagus and may be less likely to progress to dysplasia (Sharma *et al.*, 2000). A diagnosis of Barrett's oesophagus therefore should not be made histologically unless the exact site of biopsy of the metaplastic fragment is known. Barrett's oesophagus is defined as a clinicopathological diagnosis with a biopsy taken from an endoscopically visualised salmon-coloured irregularity in the lower oesophagus (Bhardwaj *et al.*, 2012).

Intestinal metaplasia can undergo inflammatory changes and becomes dysplastic. Dysplasia is assessed in columnar mucosa, and biopsies are categorised as being ‘negative for dysplasia’ if the cells show maturation towards the surface in the form of decreasing nuclear size, decreasing nuclear hyperchromasia and increasing cytoplasmic volume (Bhardwaj *et al.*, 2012). There may be histological changes deemed insufficient to characterise as dysplasia, categorised as ‘indefinite for dysplasia’. Cases as being ‘indefinite for dysplasia’ are either those with minimal or mild cytologic atypia, or those that have more than mild cytologic atypia but are accompanied by significant inflammation (Bhardwaj *et al.*, 2011). These cases need re-biopsy and evaluation.

Figure 3: Photomicrograph of Intestinal Metaplasia (H&E × 100 objective)



Barrett’s adenocarcinoma develops through a multistep process. In LGD, the molecular architecture is preserved or minimally abnormal, nuclei are elongated and crowded at the base, pseudo-stratification may be extensive, and surface villous transformation may be present (Bhardwaj *et al.*, 2011). The differentiation of LGD from non-dysplastic tissue is difficult to establish with regards to subjective histopathological criteria. High grade intestinal type dysplasia (HGD) demonstrates markedly atypical features including cytologic atypia, nuclear stratification with a loss of polarity and nuclei which are no longer radially oriented (Lomo *et al.*, 2006). OAC shows marked atypia with no radially orientated nuclei.

Figure 4: Photomicrograph of Low-Grade Dysplasia (H&E × 100 objective)

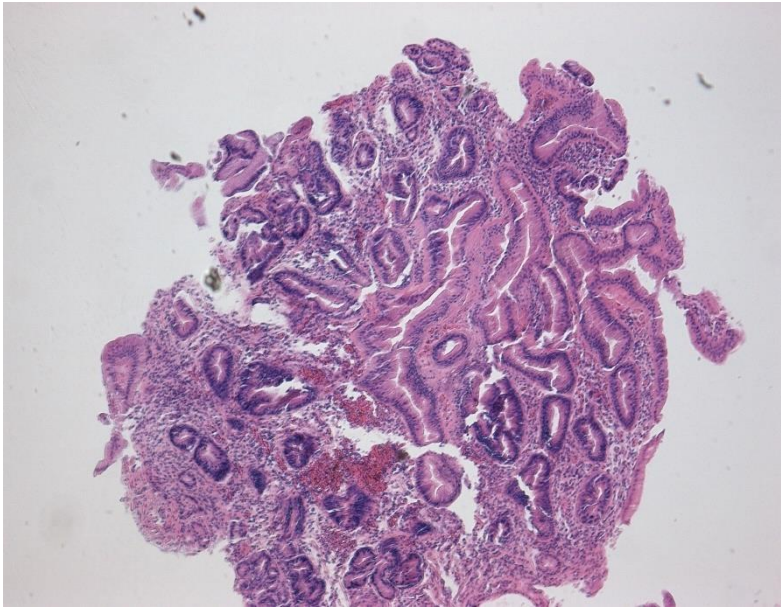


Figure 5: Photomicrograph of High-grade dysplasia (H&E × 100 objective)

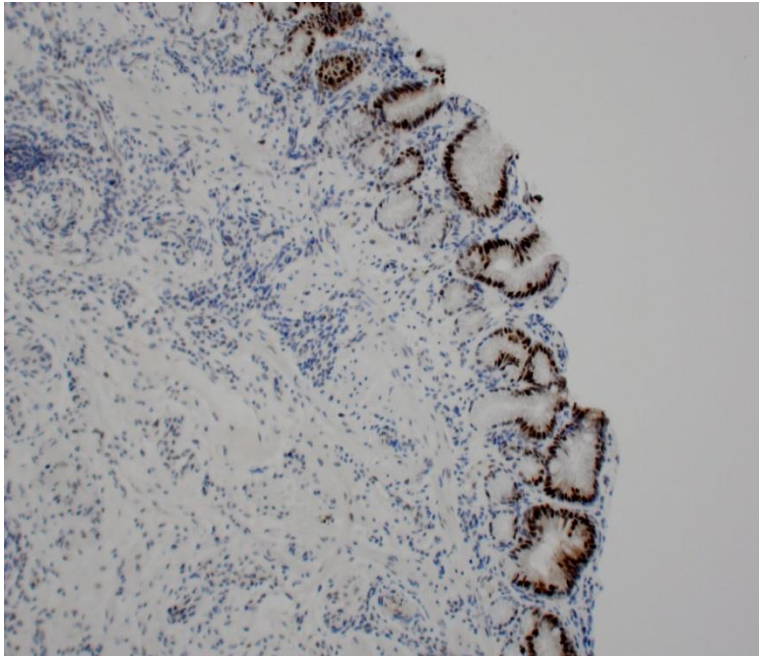
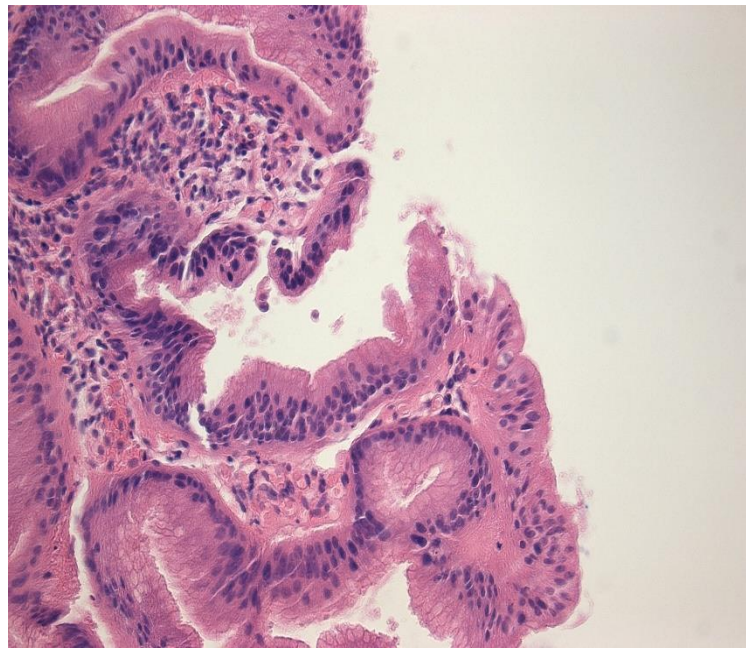


Figure 6: Photomicrograph of oesophageal adenocarcinoma (H&E × 100 objective)



Progressive histological changes are subtle hence resulting in intra and inter-observer variation in the diagnosis of dysplasia in Barrett's oesophagus. There are no defined cut-off points that distinguish disease progression when comparing LGD and HGD. Furthermore, sampling errors can occur with small dysplasia sizes and its patchy distribution. Literature from this century has compared general and specialist gastrointestinal (GI) pathologists with regards to histopathological interpretation. In a Dutch study by Kerkhof *et al.* (2007), general histopathologists were found to over diagnose HGD. Nearly 40% of patients who were initially diagnosed with HGD by a general pathologist were downgraded (11% no dysplasia, 12% indefinite for dysplasia, 16% LGD) when the samples were reviewed by three experienced GI pathologists. These results emphasise the need to obtain a second expert opinion in inconclusive cases.

Montgomery *et al.* (2001) tested criteria assessing intra- and interobserver reproducibility. The authors found that the kappa index improved from 0.43 to 0.66 once a three-tier system of grading was utilised (*i.e.*, non-dysplastic/indefinite and low grade/high grade and cancer) compared to a four-tier grading system (*i.e.*, non-dysplastic/indefinite and low grade/high grade/cancer).

Geographical discrepancies are also highlighted in the literature. In a study comparing Western and Japanese pathologists ($n = 21$ OAC lesions), 14 lesions were classified as adenocarcinoma by Western pathologists compared to 19 lesions classified as adenocarcinoma by Japanese pathologists (Schlemper *et al.*, 2000).

With large intra- and interobserver discrepancies in diagnosing dysplasia in Barrett’s oesophagus, Schlemper *et al.* (2000) proposed the ‘Vienna classification’ to reach a consensus on the nomenclature for standardisation in epithelial neoplasia. This system may be utilised to propose clear surveillance and therapeutic regimes. Although this classification system is widely used, it is qualitative and does not take disease progression into account.

Table 2: Vienna classification of epithelial neoplasia of the digestive tract (Schlemper *et al.*, 2000).

Category	Classification
1	Negative for neoplasia/dysplasia
2	Indefinite for neoplasia/dysplasia
3	Non-invasive low-grade neoplasia
	Low grade adenoma/dysplasia
4	Non-invasive high-grade neoplasia
4.1	High grade adenoma/dysplasia
4.1	Non-invasive carcinoma (carcinoma in situ)
4.1	Suspicion of invasive carcinoma
5	Invasive neoplasia
5.1	Intramucosal carcinoma
5.2	Submucosal carcinoma or beyond

Current methods utilised to reduce diagnostic error include thorough mapping protocols for all patients undergoing surveillance, independent histological inspection of specimens graded as Vienna 2–5 and expert Gastrointestinal (GI) histopathologist input for all Vienna 4–5 lesions or for lesions with persistently Vienna 2–3-graded specimens (Schlemper *et al.*, 2000). If dysplasia can initially be diagnosed accurately with adjuncts to histology using endoscopic tools or biomarkers, this would enable earlier less invasive treatment and prevent the burden of patients developing OAC.

1.1.6 Molecular and Genetic advances

Reflux injury appears to be key in the pathogenesis to OAC. The reflux of duodenogastric acid contents containing unconjugated bile acids such as deoxycholate upregulate pro-inflammatory cytokines including IL-1 β , IL-6, IL-8, and related transcription factors (notably NF κ B). These factors are known to drive metaplasia, dysplasia and cancer, and are resistant to apoptosis processes (Huo *et al.*, 2011; Fitzgerald *et al.*, 2002; O’Riordan *et al.*, 2005).

Mutations within Barrett’s oesophagus segments develop over time even in non-dysplastic epithelium. Increased clonal diversity is a marker for progression to dysplasia (Maley *et al.*, 2006). Clonal populations are stable over time, indicating that the dysplastic potential of Barrett’s oesophagus may be pre-determined. This implies that if this potential could be accurately determined at the index endoscopy, then further surveillance or management could be targeted only to those with dysplastic potential (Weaver *et al.* 2014; Dulak *et al.*, 2013).

Up to 80% of Barrett’s oesophagus is associated with *p16* anomalies, including the hypermethylation of the promoter sequence of *p16*, loss of heterozygosity (LOH), and mutation of the *p16* gene (Atherfold and Jankowski, 2006). Inactivation of *p16* results in genomic instability and uncontrolled cell division. This is in comparison with *p53* loss, which occurs later in the progression sequence towards malignancy and is not a feature of early stage non-dysplastic Barrett’s oesophagus (Reid *et al.*, 2001; Leedham *et al.*, 2008).

The prevailing view for the clonal evolution to OAC in Barrett’s oesophagus is that it occurs gradually through the sequential loss of tumour suppressors culminating in loss of *p53*. This process predominantly occurs between LGD and HGD (Kastelein *et al.* 2015). Recent literature has also suggested that *p53* mutation can lead to a more rapid progression to adenocarcinoma via chromosomal catastrophe and fragmentation (chromothripsis) or genome doubling and genetic instability (Nones *et al.*, 2014; Stachler *et al.*, 2015).

Variance in epigenetic governors disrupt normal gene expression which consequently lead to malignancy (Agarwal *et al.*, 2012) Genome studies have reported that variants in the Major Histocompatibility Complex (MHC) locus and at chromosome 16q24.1 locus, near *FOXF1*, predispose to Barrett’s oesophagus (Su *et al.*, 2012). Levine *et al.* (2013) identified three gene associations (*FOXF1*, 19p13 in *CRTC1* and 3p14 near *FOXP1*) which are implicated in oesophageal adenocarcinoma development.

In summary, Barrett's oesophagus and progression to OAC are characterised by LOH, aneuploidy, specific genetic mutations, and clonal diversity (Grady and Yu, 2018).

1.1.7 Field Effect

Numerous previous studies from the 1980's and 1990's have demonstrated multifocal high-grade dysplasia and adenocarcinoma in Barrett's oesophagus specimens, suggesting a field effect for carcinogenesis (Witt *et al.*, 1983; Cameron *et al.*, 1997). The term 'field of cancerisation' was initially used by Slaughter *et al.* (1953) when studying oral cancer. The authors summarised findings related to cancer developing in multifocal areas of precancerous change, and abnormal tissue that surrounds tumour cells.

A field effect occurs when there is an impairment of DNA damage repair mechanisms in tumour cells. This process affects the surrounding mucosa and is facilitated by adjacent inflammatory processes (He *et al.*, 2013). Reflux related changes in the distal oesophagus increase the population of regulatory T cells and activate myeloid dendritic cells (Somja *et al.*, 2013). This cytokine increase supports epithelial mesenchymal transition in the distal oesophageal mucosa thus progressing to adenocarcinoma (Underwood *et al.*, 2015).

1.1.8 Recognised Biomarkers

There is a need for reliable, robust biomarkers in the field of Barrett's oesophagus as there are no current clinical and histological criteria able to accurately predict which patients are likely to progress to HGD or OAC (Timmer *et al.*, 2013). Finding such markers would potentially reduce the number of patients required to undertake endoscopy.

The optimum marker in disease progression from Barrett's oesophagus to OAC remains an expert diagnosis of LGD (Krishnamoorthi *et al.*, 2018). Large population studies have confirmed that patients with LGD have a five times higher risk of progression to OAC compared to patients with non-dysplastic Barrett's oesophagus (Hvid-Jensen *et al.*, 2011; Bhat *et al.*, 2011). However, as a histological diagnosis of LGD is difficult due to significant inter- and intra-observer pathologist variation, multiple ongoing studies into establishing biomarkers are necessary to identify a validated prognostic tool in defining risk to neoplastic progression (Bhardwaj *et al.*, 2011).

Biomarkers in Barrett's oesophagus can be categorised into diagnostic, predictive, progression markers or prognostic biomarkers (Fouad *et al.* 2014). Diagnostic biomarkers include Trefoil Factor 3 (TFF3), a biomarker used to identify the presence or absence of specialised intestinal metaplasia. Predictive biomarkers include *p16* allelic loss which can be identified in the early stages of progression to OAC. Progression and prognostic markers include *p53*, Cyclin A and Cyclin D1. Biomarkers in Barrett's oesophagus can be also be classified related to DNA abnormalities and panel profiling (Table 3).

Immunohistochemical (IHC) biomarkers have been interrogated in Barrett's oesophagus and its progression to OAC as they can be applied to standard histological samples. Early IHC studies of nuclear *p53* expression in patients with Barrett's oesophagus have shown to improve inter-observer variability in diagnosing dysplasia and can predict progression risk with an odds ratio (OR) of 3–8 (Skacel *et al.*, 2002; Kaye *et al.*, 2010). Despite concerns with regards to the reproducibility of this assay (*p53* positivity rate in Barrett's dysplasia ranges from 50% to 89% - Khan *et al.*, 1998), the BSG has proposed that the addition of *p53* immunostaining to the histopathological assessment process may improve the diagnostic reproducibility of a diagnosis of dysplasia in Barrett's oesophagus (Fitzgerald *et al.*, 2014).

Janmaat *et al.* (2017) conducted a systematic review and meta-analysis interrogating 16 different biomarkers in 36 studies. The authors found that aberrant *p53* expression was significantly associated with an increased risk of neoplastic progression with an OR of 3.18 (95% CI 1.68 to 6.03). This association was confirmed for both non-dysplastic Barrett's oesophagus and Barrett's oesophagus with LGD (Janmaat *et al.*, 2017).

Table 3: Summary of molecular biomarkers (Timmer *et al.*, 2013; Fouad *et al.*, 2014)

Biomarker	Sample size	Baseline histology	Endpoint
DNA abnormalities			
Aneuploidy / Tetraploidy	322	SIM, Indefinite, LGD	OAC
Biomarker Panels			
LOH	243	SIM	OAC
Expert LGD, <i>Aspergillus oryzae</i> lectin	380	SIM, Indefinite, LGD	OAC
8 gene methylation	195	SIM	HGD / OAC
Epigenetics			
<i>p16</i> methylation	53	SIM, LGD	HGD/ OAC
Tumour suppressor loci			
<i>p53</i> staining	48	LGD	HGD / OAC
<i>p53</i> LOH	256	SIM, Indefinite, LGD	OAC
Cell Cycle markers			
Cyclin A	48	SIM	HGD / OAC
Cyclin D1	307	SIM	OAC
Clonal diversity			
Clonal diversity measures	239	SIM	OAC
Proliferation			
Mcm2	27	SIM	OAC
Serum Biomarkers			
Selenoprotein P	361	Variable	OAC
Leukocyte telomere length	300	Variable	OAC

A recent Dutch study from 2019 interrogated 130 patients who progressed from Barrett's oesophagus to OAC vs. 130 patients who did not progress from Barrett's oesophagus (Duits *et al.*, 2019). The authors concluded that this case-control study demonstrated an expert LGD diagnosis (OR 35.7), abnormal *p53* expression (OR 4.1) and abnormal expression of *Aspergillus oryzae* lectin (OR 4.3) were all independently associated with progression to HGD/OAC.

The ideal biomarker has to be cost-effective, minimally invasive and superior to current diagnostics involving endoscopy and multiple biopsies. There is a lack of prospective controlled trials in this field attributable to high costs related to specimen collection and storage. No clear data is available supporting the use of markers which can sub-select those at higher risk of progression, other than an expert diagnosis of LGD (Stone *et al.*, 2004). Further multicentre research is necessary to aid in prompt LGD diagnosis in these patients. This could subsequently reduce the burden of OAC developing in the future if less invasive regimes are utilised at this point.

1.1.9 Screening

Screening is based on the presence of multiple risk factors including chronic GORD, male sex, white race, patients > 50 years and a high BMI (Gordon *et al.*, 2014; Wang *et al.*, 2008). The decision on who to be screened should however be individualised. A large number of asymptomatic patients may miss the opportunity for their cancers to be detected early. Lagergren *et al.* (1999) found that up to 40% of patients with OAC had no history of chronic GORD. Furthermore, sampling and diagnostic errors with inter-variable pathological discrepancies results in a reduced effectiveness of screening (Montgomery *et al.*, 2001).

The gold standard method of screening is visual OGD inspection and four quadrant biopsies of mucosal irregularities in salmon-coloured mucosa above the GOJ at every 1–2 cm intervals using the Seattle protocol (Sampliner, 1998). This has been clarified by NICE guidelines on management of dyspepsia published in 2014 (NICE 2014, accessed 07/11/2019). They recommend that OGD should be considered if a patient with GORD has risk factors including older age, male gender and a history of reflux. Other risk factors include a long duration of symptoms, an increased frequency of symptoms, previous oesophagitis, previous hiatus hernia, and evidence of oesophageal stricture or oesophageal ulcers.

Saad *et al.* (2003) established that standard brush cytology demonstrated a high diagnostic sensitivity for HGD/OAC (sensitivity 90% vs. 94%), moderate sensitivity for Barrett's oesophagus (60% vs. 92%) and low sensitivity for LGD (20% vs. 97%) compared with histology. Alexander *et al.* (1997) commented that although brush cytology compliments histology there is an added increased cost with no true improvement.

Standard OGD is expensive and associated with a small risk of complications such as bleeding, perforation, aspiration and cardiopulmonary events (Sanghi *et al.*, 2019). Since OGD is not a suitable method for screening in large populations, there is a need for an alternative, cheap, widely available and an accurate method of screening (Offman *et al.*, 2017).

The ACG has recommended unsedated trans-nasal endoscopy as an alternative to traditional endoscopy for screening in Barrett's oesophagus. Shariff *et al.* (2012) reported that with topical anaesthesia, unsedated transnasal endoscopy was safer with fewer procedure- and sedation-related complications compared to standard OGD. Jobe *et al.* (2006) established the sensitivity and specificity for the detection of specialised intestinal metaplasia was 91% and 100% with unsedated transnasal endoscopy compared to standard OGD.

Non-invasive methods such as the cytosponge or capsular endoscopy have been utilised to screen for specialised intestinal metaplasia but do not alter the difficulty of sub-selecting a population with an increased prevalence of specialised intestinal metaplasia or OAC. Cytosponge is a mesh surrounded by a gelatin capsule attached to a string which is passed transorally (Ross-Innes *et al.*, 2015). The capsule dissolves in the proximal stomach 5 minutes post ingestion, expanding the mesh to a sphere. The cytological specimen is then stained with Trefoil Factor 3 (TFF3) which is a biomarker for specialised intestinal metaplasia.

Kadri *et al.* (2010) found that the cytosponge with TFF3 had a sensitivity of 73.3% (95% CI 44.9–92.2%) and a specificity of 93.8% (95% CI 91.3–95.8%) for detecting Barrett’s oesophagus ≥ 1 cm of circumferential length. Heberle *et al.* (2017) carried out a cost-analysis and established that screening GORD patients with cytosponge and following up positive results with OGD for confirmation reduced cost by 27–29% when compared with screening by OGD alone.

Capsular endoscopy allows oesophageal visualisation using wireless cameras without obtaining biopsies. Capsular endoscopy has reported a sensitivity and specificity of 77% and 86% for a diagnosis of Barrett’s oesophagus, respectively, compared to standard OGD, but just 73% specificity compared with histologically confirmed specialised intestinal metaplasia in a meta-analysis of 9 studies ($n = 618$) (Bhardwaj *et al.*, 2009).

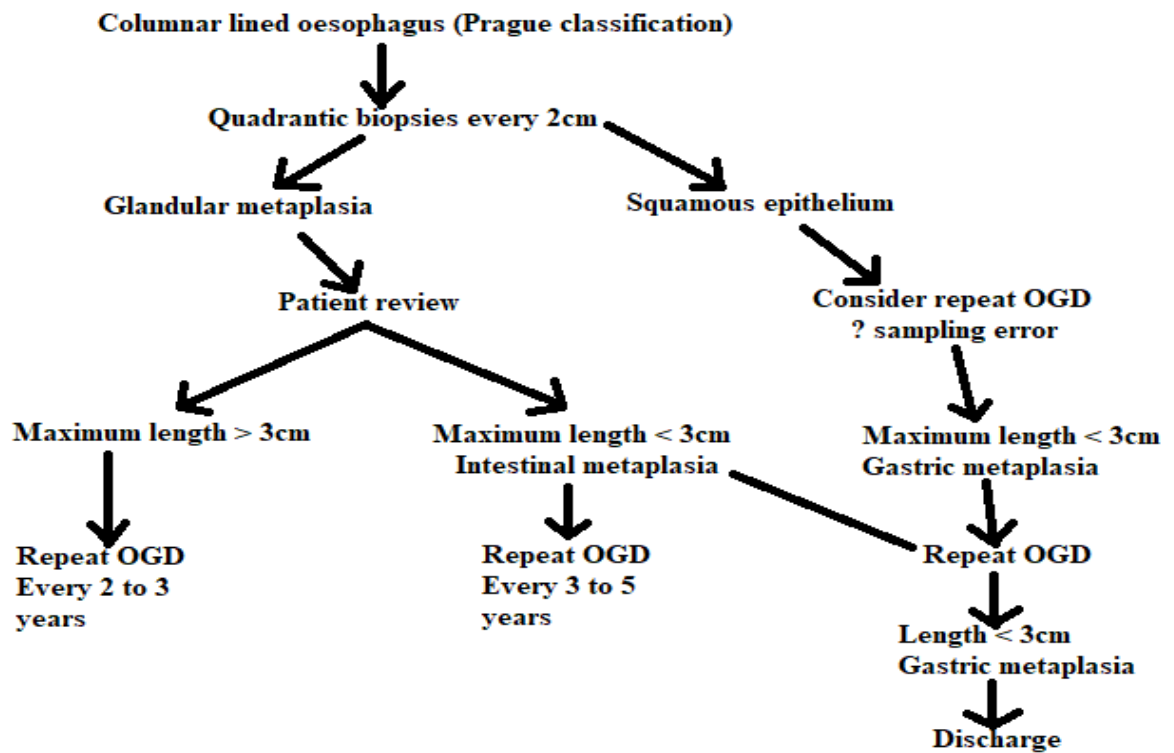
Liquid biopsies utilising blood samples and extracting circulating microRNAs expressed in disease are gaining promise as a screening tool. Bus *et al.* (2016) profiled circulating microRNAs in patients with Barrett’s oesophagus. The authors found that in 41 patients with Barrett’s oesophagus and 15 controls, a panel of 4 circulating miRNAs (miRNA-95-3p, -136-5p, -194-5p, and -451a) distinguished Barrett’s oesophagus from controls with a sensitivity and specificity of 78% and 86%, respectively.

1.1.10 Surveillance

The primary aim of surveillance in Barrett's oesophagus is to detect dysplasia and malignancy before distant disease has occurred. OAC usually presents with advanced disease as a result of early lympho-vascular submucosal invasion (Dunbar *et al.*, 2012). OGD remains the primary tool for surveillance using the Seattle protocol (Provenzale *et al.*, 1999). The frequency of surveillance is determined by the degree of dysplasia at biopsy. Biopsies are then classified as per the Vienna classification (Schlemper *et al.*, 2000). It is important to note that surveillance endoscopy should be performed in patients whose reflux symptoms are controlled, reducing the probability of reactive changes interfering with pathological interpretation (Hanna *et al.*, 2006).

In the United Kingdom, new guidelines with regards to surveillance and management were published in 2014 by the British Society of Gastroenterologists (BSG). The authors performed a systematic review of the literature published up until December 2012 in order to address controversial issues in Barrett's oesophagus including definition, screening and diagnosis, surveillance, pathological grading for dysplasia, management of dysplasia, and early cancer.

Figure 7: Surveillance flow chart for non-dysplastic Barrett's oesophagus



For metaplasia only, surveillance every 2-5 years is offered to patients. Patients should be given information regarding the risks and benefits of surveillance. Worldwide guidelines differ with regards to differing Barrett's oesophagus segment length. The BSG guidelines state that endoscopy should be repeated 3-5 years if the maximal length is < 3 cm, and every 2-3 years if above or equal to 3 cm (Fitzgerald *et al.*, 2014).

Evidence for improved outcomes from surveillance is weak and remains the subject of debate. The UK multicentre BOSS (Barrett's Oesophagus Surveillance Study) trial aims to compare the benefits of 2-yearly surveillance endoscopy against endoscopy on an 'at need' basis only. The primary endpoint is overall survival. This randomised trial will provide data to evaluate the efficacy and cost-effectiveness of screening Barrett's oesophagus patients for OAC.

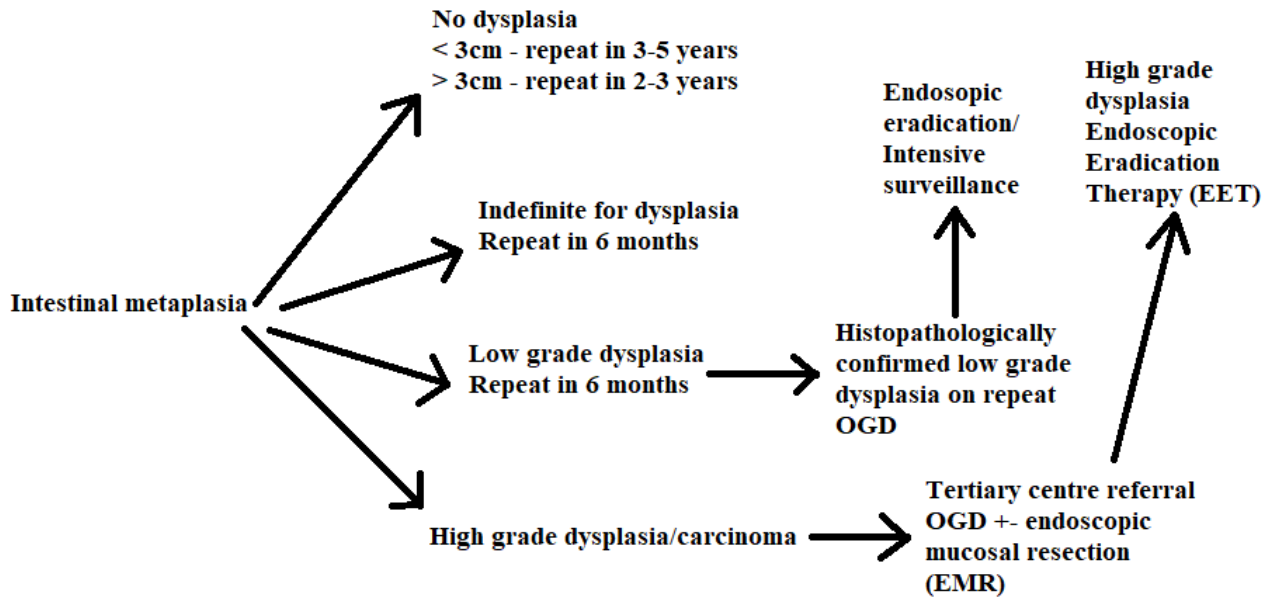
Some literature does demonstrate a survival advantage in patients with Barrett's oesophagus undergoing surveillance. El Serag *et al.* (2016) found that patients diagnosed with OAC during surveillance were detected at an earlier stage (stage 0 to 1: 74.7% vs. 56.2; $p < 0.001$), survived longer (median 3.2 vs. 2.3 years; $p < 0.001$), and had lower cancer-related mortality (34.0% vs. 54.0%, $p < 0.0001$) compared with those not in surveillance. As the natural course of Barrett's oesophagus is unknown, and surveillance is expensive and time-consuming, surveillance has been the subject of much scrutiny (Saxena *et al.*, 2017).

A cost-effective analysis of surveillance in patients with non-dysplastic disease at 5 yearly intervals was found not to be cost effective (Kastelein *et al.*, 2015). The authors found that unless the annual progression rate to adenocarcinoma were 1.9% then a QALY (Quality Adjusted Life Year) threshold of <€50,000 could not be achieved (Kastelein *et al.*, 2015). This finding is similar to previous studies highlighting that surveillance in this group is not cost-effective (Provenzale *et al.*, 1999).

Findings such as inflammation and ulceration considered to be indefinite for dysplasia may evolve as a result of erosive oesophagitis. Acid suppression should be offered and a repeat endoscopy should be offered in 6 months (Fitzgerald *et al.*, 2014).

If a histological finding of LGD is encountered, a repeat endoscopy should be performed at 6 months. If this confirms the diagnosis, patients should be offered endoscopic surveillance (every 6 months for 2 years, annually thereafter) or eradication therapy (radiofrequency ablation, endoscopic resection, photodynamic therapy, spray cryotherapy). Given the observation that the progression from LGD to adenocarcinoma is low (approx. 5 per 1000 patient years), the absolute benefit of eradication therapy for LGD is not certain (de Jonge *et al.*, 2010).

Figure 8: Surveillance and management for dysplasia (LGD/HGD)



HGD should be confirmed by two expert GI pathologists and these patients should be referred to a tertiary centre for consideration of repeat endoscopy with biopsies, endoscopic mucosal resection and/or eradication therapy.

1.1.11 Management

All patients with non-dysplastic Barrett's oesophagus should be on acid suppression therapy. However, there is no convincing evidence that this reverses specialised intestinal metaplasia (Spechler *et al.*, 2011). LGD should be managed with endoscopic surveillance biopsies with endoscopic resection (Wang *et al.*, 2008). Endoscopic radiofrequency ablation (RFA) has been utilised in LGD patients. An American multicentre retrospective cohort study discovered a 0.8% progression rate amongst patients with LGD diagnosed by expert pathologists and treated with RFA compared with 6.6% in the control surveillance group (Small *et al.*, 2015). No strong data is available supporting the use of markers which can sub-select those at higher risk of progression, other than an expert diagnosis of LGD (Krishnamoorthi *et al.*, 2018). Patients with multifocal areas of dysplasia may have an increased risk of progression, as may patients with persistent LGD over time (Srivastava *et al.*, 2007).

Index ablation has been shown to be a cost-effective strategy compared to intensive surveillance. A cost-effectiveness analysis study stated that ablation with RFA in patients with LGD is more cost-effective than surveillance if ablation permanently eradicated LGD in more than 28% of patients, without the need for further surveillance in this group (Inadomi *et al.*, 2003).

Previous BSG and AGA guidelines did not recommend ablation for LGD. The SURF (Surveillance *vs.* Radio Frequency ablation) trial randomised patients with LGD to RFA ($n = 68$) *vs.* intensive surveillance ($n = 68$). Results published in 2014 found that ablation reduced the risk of progression to HGD or adenocarcinoma from 26.5% in the control arm to 1.5% in the RFA arm (95% CI 14.1–35.9%, $p > 0.001$). When RFA was utilised, 88.2% of intestinal metaplasia was eradicated and 92.6% of dysplasia using RFA, *vs.* rates of 0 and 27.9% in the surveillance arm (Phoa *et al.*, 2014). Since the publication of these findings, up to date ACG and BSG guidelines recommend endoscopic eradication with RFA as a recognised treatment of choice for LGD although endoscopic surveillance is regarded as a valid alternative (Spechler *et al.*, 2011).

Patients with HGD should be given options including endoscopic therapies including RFA +- endoscopic mucosal resection, surgical resection, or intensive surveillance. There should be discussions involving the patient's fitness for surgery, the patient's desires and the expertise at a particular centre.

Ablative therapies are able to treat entire Barrett's oesophagus segments. Surgery should only be necessary for patients with risk factors for lymph node metastases. Current standards reserve oesophagectomy for patients with T1b invasion (submucosal invasion), multifocal carcinoma or lesions not amenable to endoscopic resection (Buttar *et al.*, 2001).

A recent cost-effectiveness analysis was conducted for endoscopic eradication therapy (EET) for treatment of all grades of dysplasia in Barrett's oesophagus patients (Pollit *et al.*, 2019). Endoscopic eradication therapy for patients with LGD and HGD arising in Barrett's oesophagus is cost-effective compared to endoscopic surveillance alone (lifetime £3,006 per QALY gained). The authors further concluded that as the time elapses, the treatment becomes more cost-effective. The five-year financial impact to the UK NHS of introducing EET was estimated as £7.1m (Pollit *et al.*, 2019; Booth *et al.*, 2012).

Photodynamic treatment (PDT) using 5-aminolevulinic acid and porfimer sodium has been shown to be inferior to RFA. In a multicentre study, 13% of patients progressed to adenocarcinoma despite treatment (Overholt *et al.*, 2007) and no studies with long-term follow-up have shown improved overall survival vs. oesophagectomy (Prasad *et al.*, 2007).

The current guidelines state that patients should continue ongoing surveillance. There is no long-term data (>5 years) on the recurrence of intestinal metaplasia or dysplastic changes in squamous epithelium. Following eradication, surveillance should be carried out at 3-monthly intervals for 1 year and either 6-monthly for the next year and annually thereafter (AGA) or annually after the first year (ACG).

1.1.12 Summary

Barrett's oesophagus is the only known precursor to OAC with a population prevalence of around 1 – 2% (Zagari *et al.*, 2008). Established risk factors include older age, male gender and a history of reflux symptoms (Gordon *et al.*, 2014). Although guidelines on the screening and surveillance exist in Barrett's oesophagus, the current strategies are inadequate as more than 90% of patients diagnosed with OAC do not have a prior diagnosis of Barrett's oesophagus (Graham *et al.*, 2016). Furthermore, the annual risk for developing OAC has been shown in large population studies to be as low as 0.16% (Bhat *et al.*, 2011).

The ACG has suggested unsedated transnasal endoscopy as an alternative to conventional upper endoscopy in patients with risk factors for Barrett's oesophagus as a screening tool (Sanghi *et al.*, 2019). The BSG currently do not suggest OGD as a feasible screening test for an unselected population with gastro-oesophageal reflux symptoms (Fitzgerald *et al.*, 2014). Standard OGD is the gold standard method in surveillance in Barrett's oesophagus. NICE recommends considering OGD surveillance to check for OAC progression in patients whom already have a diagnosis of Barrett's oesophagus confirmed by endoscopy and histopathology. The BSG currently do not suggest OGD as a feasible screening test for Barrett' oesophagus.

The need for an alternative, non-invasive method of screening and/or surveillance could be highly beneficial. This could subsequently reduce waiting times, alleviate patient anxiety and reduce future costs related to multiple invasive procedures and management in the NHS.

1.2 Advanced imaging technology in Barrett's Oesophagus

Thorough endoscopic assessment and biopsies are key to a diagnosis of Barrett's oesophagus and subsequent surveillance. Most gastroenterologists and upper GI surgeons appreciate quadrant biopsies are time-consuming and advanced imaging would be a helpful additional tool.

1.2.1 High resolution endoscopy

High-resolution white light endoscopy enhances mucosal visualisation combining pixelated endoscopes (up to 1,000,000) with high definition screens. High-resolution endoscopy has demonstrated a greater sensitivity in the detection of early neoplastic lesions when compared to standard endoscopy (Kara *et al.*, 2005 (A)).

1.2.2 Chromoendoscopy

Chromoendoscopy is a diagnostic tool where a chemical stain is sprayed onto the mucosal surface to highlight specific regions of interest. Stains can be categorised as 'absorptive' (acetic acid, methylene blue, lugol solution) and 'non-absorptive' (indigo carmine). Advanced imaging modalities magnify the view and subsequently increase the probability of finding suspicious lesions. Studies have demonstrated an increased diagnostic yield using chromoendoscopy in recognising dysplasia in Barrett's oesophagus compared to random biopsies (Pohl *et al.*, 2010).

Acetic acid staining with targeted biopsies has shown an increased detection rate of dysplasia and OAC, even with white light standard endoscopy (Fortun *et al.*, 2006). Longcroft-Wheaton *et al.* (2010) has demonstrated that acetic acid chromoendoscopy yielded a 95.5% sensitivity and 80% specificity for the detection of OAC.

Indigo carmine used as a non-absorptive contrast stain has shown to be highly sensitive (83%) and specific (88%) for HGD (Sharma *et al.*, 2006 (B)). The agent is currently not able to differentiate between specialised intestinal metaplasia and dysplasia (Sharma *et al.*, 2003).

1.2.3 Autofluorescence

Autofluorescence imaging utilises short wavelengths of light to stimulate endogenous substances (nicotinamide adenine dinucleotide (NADH), collagen, aromatic amino acids and porphyrines) in tissue to emit fluorescent light of a longer wavelength (Trivedi *et al.*, 2013). This interrogates the tissue at depth and aids in the interpretation of vasculature and topography (DaCosta *et al.*, 2006). Studies have not consistently demonstrated the superiority of this method compared with high resolution endoscopy for the detection of dysplasia (Kara *et al.*, 2005 (B); Borovicka *et al.*, 2006). Consequently, autofluorescence guided biopsies are not currently employed in the NHS.

1.2.4 Narrow band imaging (NBI)

Narrow-band imaging allows visualisation of the superficial mucosa and vasculature without the need for additional dyes. NBI illuminates mucosa with blue and green wavelength light thus demonstrating tissue vasculature.

A meta-analysis by Mannath *et al.* (2010) evaluated 446 patients. The authors found pooled sensitivity and specificity were 96% (95% confidence interval [CI] 0.93-0.99) and 94% (95 % CI 0.84-1.0) on a per-lesion analysis with similar results on a per-patient analysis for diagnosing HGD. For specialised intestinal metaplasia characterisation, the pooled sensitivity and specificity were 95% (95% CI 0.87-1.0) and 65% (95% CI 0.52-0.78) on a per-lesion analysis. The authors concluded that magnified NBI is accurate with high diagnostic precision for diagnosis of HGD in Barrett's oesophagus. They further commented that NBI has high sensitivity but poor specificity for characterising specialised intestinal metaplasia.

1.2.5 Optical coherence tomography (OCT)

Optical coherence tomography (OCT) utilises electromagnetic (EM) waves to generate images based on the detection of reflected light. Resolutions up to 10 - 25 μm enables the identification of microscopic features such as lymphovascular structures (DaCosta *et al.*, 2003). Robles *et al.* (2015) interrogated 19 studies (17 *in vivo*; 2 *ex vivo*). The authors found an excellent diagnostic yield for specialised intestinal metaplasia detection but a poor yield for dysplasia. Evans *et al.* (2006) only demonstrated an 83% sensitivity and 75% specificity between differentiating HGD and OAC using OCT.

1.2.6 Confocal fluorescence microendoscopy

Confocal fluorescence microendoscopy images fluorophores within the cell microstructure and generates a histological image (Kara *et al.*, 2007). Kara *et al.* (2006) evaluated 63 patients in *ex vivo* samples. The authors concluded that a differentiation could be made between Barrett's oesophagus and HGD, but that dysplastic tissue evaluation needed histological guidance.

Curvers *et al.* (2008) concluded that the above enhanced imaging techniques may be no better than using high-resolution white light endoscopy. Furthermore, these methods fail to achieve the aim of replacing random biopsies for diagnosis. This encourages the need for innovative, diagnostic screening and surveillance modalities to be explored in the field of Barrett's oesophagus and its subsequent transformation to OAC.

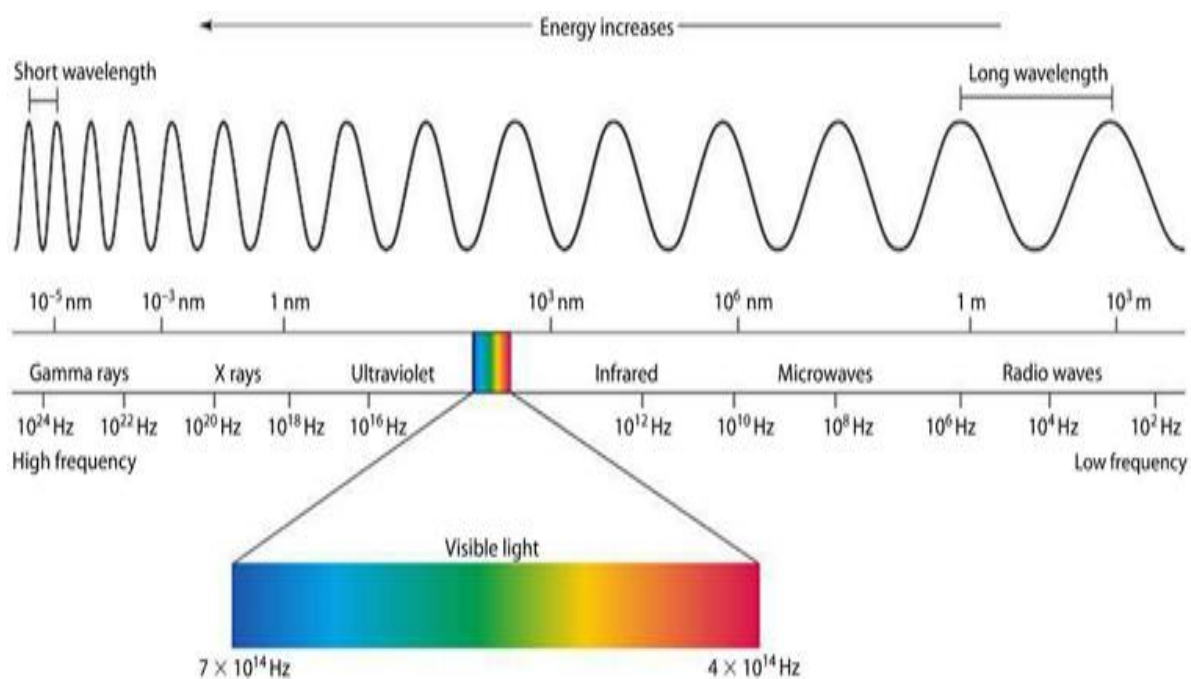
1.3 Vibrational spectroscopy

1.3.1 Principles

The electromagnetic (EM) spectrum is the range of frequencies of electromagnetic radiation and their respective wavelengths and photon energies. The electromagnetic spectrum covers frequencies ranging from below one hertz to above 10^{25} hertz. Lower frequency spectra have longer wavelengths (*e.g.*, radiowaves) and higher frequency spectra have shorter wavelengths (*e.g.*, gamma rays).

Figure 9: Electromagnetic radiation spectrum

(<https://leadertechinc.com/blog/wp-content/uploads/2016/08/Electromagnetic-Spectrum.png>) accessed 20/08/2019



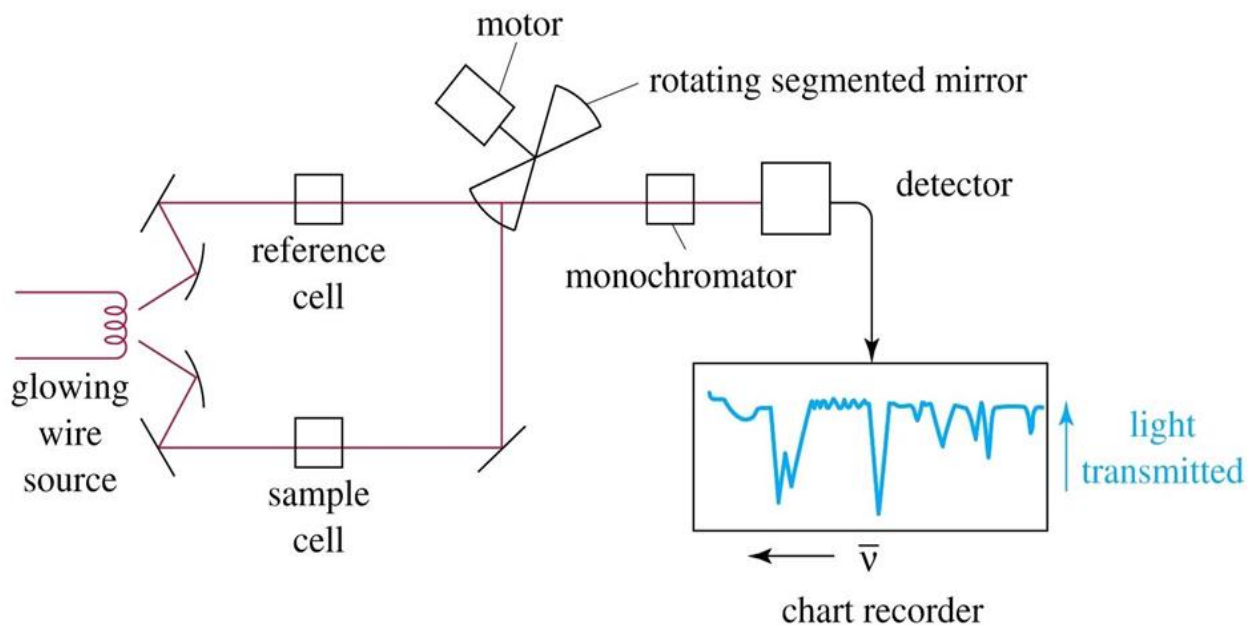
1.3.2 Attenuated total reflection (ATR) / Fourier-transform infra-red spectroscopy (FTIR)

Spectroscopy offers the ability to detect subtle biochemical changes in tissues without destroying it. Subsequent spectra is based on the interaction between light and the tissue. This spectrum provides key information at the micro-cellular level, including data on DNA, lipids, proteins and carbohydrates.

Fourier-transform infrared spectroscopy (FTIR) is an analytical tool which aids in the characterisation and identification of organic molecules. In Attenuated Total Reflection (ATR) mode, this type of spectroscopy enables functional groups present over a depth of about $1\ \mu\text{m}$ to be identified. Infrared absorption occurs if the energy of an incident photon from a polychromatic light source matches the energy gap between the ground state of a molecule and an excited vibrational state. For simple vibrations within molecules, the matching frequency range is in the mid-range infrared region ($400\ \text{cm}^{-1} - 4000\ \text{cm}^{-1}$). This corresponds to wavelengths of about $10\ \mu\text{m}$. Chemical bonds in the media vibrate symmetrically, non-symmetrically, bend, or vibrate along their x or y axes. This flux results in light formation. The energy absorbed at each wavelength can be established by measuring this light.

Figure 10: FTIR spectrometer diagram

(<http://cssmith.co/wp-content/uploads/2017/11/ftir-spectrometer-diagram-labeled.jpg>) accessed 20/08/2019



During analysis, the sample is kept in contact with a crystal allowing total internal reflection. An infrared ray arrives at the crystal where the material under study has been placed. The internal reflection of the ray gives rise to an evanescent wave which, at each reflection, continues beyond the surface of the crystal and penetrates the sample. The penetration depth depends on the wavelength, the angle of incidence of the beam on the crystal and the nature of the crystal. Spectra are thus obtained (curves of absorbance vs. wavelength) that have absorption peaks characteristic of the functions present at the surface of the medium.

Data can be collected quickly from analysing a small volume of fluid or tissue without damaging it. The process however has limited uses *in vivo* as biological tissue has a high-water content, therefore adding a complexity in analysis. Water is absorbent at the mid IR range of 400 cm^{-1} to 4000 cm^{-1} . This can mask vibrations from the medium itself.

1.3.3 Raman Spectroscopy (RS)

The phenomenon of Raman scattering of light was first observed experimentally in 1928 by C.V. Raman. The majority of light is scattered without any interaction of photons with the material being studied. This is regarded as elastic (Rayleigh) scattering. In Raman spectroscopy, spectral measurements can be made on molecular material based on inelastic scattering of monochromatic radiation from a laser in the visible, near infrared, or near ultraviolet range. Energy is exchanged between the photon and the molecular material. The resultant scattered photon is of higher or lower energy than the incident photon.

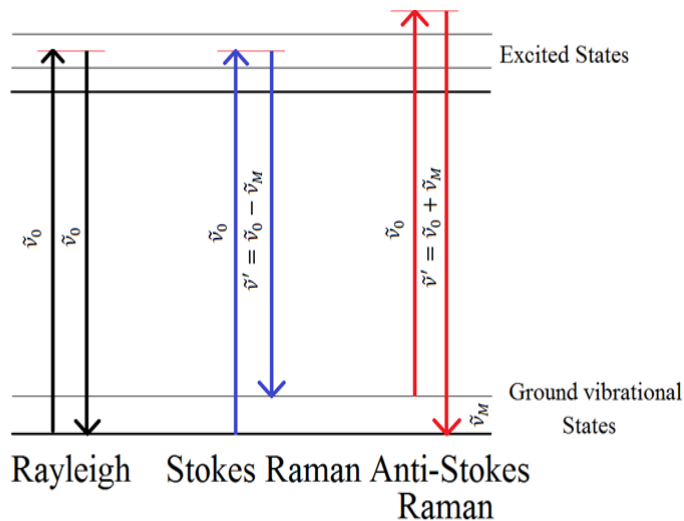
The Raman effect states that a very small proportion of the incident photons are scattered inelastically (Stokes- and Anti-Stokes Lines). The energy difference corresponds to the energy change of the molecule, which refers to the transition between two vibrational states.

A Stokes shift occurs when the inelastically scattered photon is shifted to a lower frequency (energy level higher than original state). An anti-stokes shift occurs when the photon will be of higher frequency. The scattered spectra obtained can give qualitative information regarding the chemical composition in DNA, RNA and other molecules, as well as quantitative information as the intensity of scattering is proportional to the concentration of biomolecules in the medium.

Raman spectroscopy analyses samples with speed and precision and has enormous potential for the detection of biochemical changes at a molecular level. Raman spectroscopy measures relative frequencies at which a sample scatters radiation, unlike IR spectroscopy which measures absolute frequencies at which a sample absorbs radiation.

Figure 11: Rayleigh, Stokes and Anti-stokes scattering

<https://chem.libretexts.org/@api/deki/files/9660/Raman.png> (accessed 30/08/2019)



1.3.4 Properties of biofluids and tissue in Spectroscopy

The advent of biospectroscopy coincides with the need for safe, rapid, non-destructive, real-time techniques of organic and inorganic compound analysis in human fluids and tissues. This would allow biochemical characterisation improving the response time to clinical diagnoses and alleviating patient anxiety after initial clinical consultations.

1.3.4.1 Biofluid analysis

Biofluid analysis has been a favourable technique in vibrational spectroscopy for its high-throughput. Measurement conditions can be controlled, the required sample volume is small and the signal can be enhanced through non-linear optical effects and/or metallic nanoparticles (Kneipp *et al.*, 2002; Aroca, 2004; Chen *et al.*, 2012). The major limitations are of measurement variations due to post-collection procedures such as sample dilution, storage at cold temperatures, drying effects and the addition of anticoagulants (Leal *et al.*, 2018). Multicentre robust studies are necessary in order to translate this into the clinical setting (Lovergne *et al.*, 2016).

Whole blood is a mixture of cellular elements, crystalloids and colloids. Its function in mammals is transporting electrolytes, oxygen, nutrients and products of cellular metabolism. Over the last two decades, serum has predominantly been used in studies in spectroscopy. The reason for this is the reduced cost compared to plasma separation and segregation as anticoagulants affect IR and Raman spectra (Rohleder *et al.*, 2005). In the NHS, blood analysis is a routine, commonly utilised method for laboratory analysis contributing to clinical diagnoses. Its analysis allows the identification and differentiation of hematopoietic cells and can biochemically quantify the compounds requested by physicians to evaluate the patient's clinical condition (Managò *et al.*, 2018).

Saliva is a biological fluid containing water, proteins (peptides) and inorganic substances. Its function acts to maintain a moist oral mucosa which is less susceptible to abrasion, and removal of micro-organisms, desquamated epithelial cells, leucocytes and food debris by swallowing (Dawes *et al.*, 2015). Salivary analysis has been utilised in the diagnosis and the monitoring of treatment for conditions including Sjögren's syndrome, cystic fibrosis and sarcoidosis. Its advantages of use include its simple and non-invasive collection, as well as the small volumes necessary for analysis. However, the biological variability of saliva is great, depending on factors such as smoking and circadian rhythms (Diem *et al.*, 2016). This issue could be resolved by standardising sample collection (spitting vs draining method) and its stimulation of production (with or without exogenous stimuli) (Leal *et al.*, 2018).

Urinalysis has been utilised to delineate quantification of urea, creatinine, excess excretion of protein, excess excretion of microalbumin and glomerular filtration rate (GFR) (Leal *et al.*, 2018). In addition, 24-hr protein dosage, serum urea, and creatinine content were associated with the concentration of these compounds in blood (Leal *et al.*, 2018). Urine is used routinely by clinicians and researchers and has clear advantages of non-invasive collection and easy storage. Urinalysis can however be complex depending on patient factors such as BMI as well as differences in institutional reference values making data comparison complex.

There appears to be expected non-modifiable biological factors affecting biofluid analysis, such as advancing age, biological rhythms and physiological changes. Many laboratory parameters vary depending on the time of day, week, month or year when they are sampled. Body temperature, hormone production (e.g. cortisol, testosterone) and cardiac function follow a circadian (24 hour) rhythm. To allow for this effect, some laboratory tests are recommended at specific times of the day, *e.g.*, testosterone should be sampled between 7 am – 10 am. There have been no studies performed assessing IR and Raman spectral discrepancies with biofluids related to diurnal variation.

A patient's diet, comorbidities and lifestyle factors can all have a pre-analytical influence on laboratory parameters. Malnutrition has varying effects on laboratory results, depending on the nature of the patient's nutritional status. Nicotine exposure with smoking can have acute and chronic effects on laboratory investigations. The mechanisms behind these changes are not fully understood (Guder *et al.*, 2003). Finally, medications can have both direct and indirect effects on biomolecular structure.

It is unclear as to how these individual factors would clearly affect IR and Raman spectral data in biofluids. This would warrant substantial further investigation. Standard operating protocols have been utilised in the pre-processing stage of biofluid analysis with vibrational spectroscopy. Non-modifiable biological factors and patient lifestyle factors would need to be taken into account when performing future spectroscopic analysis.

1.3.4.2 Human tissue analysis

The morphological classification of certain tumours is becoming difficult even with the advent of staining and other histopathological adjuncts. Recent literature suggests that vibrational spectroscopy has molecular sensitivity toward biochemical changes in tissue rivalling IHC (Diem *et al.*, 2016). Vibrational spectroscopy detects changes in the metabolome and proteome. The techniques of ATR-FTIR and Raman are quick, cheap and require little tissue material.

1.4 The role of Biospectroscopy in cancer diagnostics

The use of biofluids in biospectroscopy for disease diagnosis and screening is possible in a wide range of conditions including cancer. Further substantial prospective trials are necessary to delineate whether biospectroscopy has the ability to identify the small number of at-risk individuals amongst the large number not requiring follow up (Mitchell *et al.*, 2014).

Over the past 20 years there has been a large shift towards biofluid sampling to identify biomarkers of malignancies including the prostate, brain, of gynaecological origin and the GI tract. Studies from as early as 1983 detected that serum DNA concentrations are markedly elevated in malignancy, and moderately elevated in benign disease, as compared with normal controls (Shapiro *et al.*, 1983).

ATR-FTIR spectroscopy has shown promise in biofluid analysis for ovarian cancer. Gajjar *et al.* (2013) performed a plasma analysis of patients with ovarian cancer. The authors found that approximately 97% of patients were correctly diagnosed when compared to gold standard histological diagnosis. Paraskevaidi *et al.* (2018) used ATR-FTIR to analyse urine samples from women with endometrial and ovarian cancer, as well as from healthy individuals. The authors found high levels of accuracy for both endometrial (95% sensitivity, 100% specificity) and ovarian cancer (100% sensitivity, 96.3% specificity). ATR-FTIR spectroscopy has also been used in conjunction with specific serum assays to increase the diagnostic accuracy and detection of intracranial gliomas (Hands *et al.*, 2014).

Raman spectroscopy of biofluids has been conducted in the spheres of gynaecology and gastroenterology (Paraskevaidi *et al.*, 2019). Raman spectroscopy has shown encouraging results in testing biofluids in cirrhosis patients differentiating between patients with or without hepatocellular carcinoma (HCC). Spectrochemical analysis has identified small spectral differences with an accuracy of approximately 91% (Taleb *et al.*, 2013).

There are ongoing studies and research both in the UK and USA with regards to biomarkers in biofluids. The lack of specificity of CA-125 in ovarian malignancy limits its role as an effective screening test. Many cancers have no specific marker. Finding one such biomarker would allow imaging free follow up, thus reducing subsequent radiation exposure in patients undergoing yearly Computerised Tomography (CT) imaging. In the sphere of Barrett's oesophagus, dysplasia and OAC disease classification, there have been no substantial studies looking into ATR-FTIR and Raman spectroscopic analysis of plasma, serum, urine or saliva to classify disease to OAC.

Vibrational spectroscopic techniques have been used to delineate classification of oesophageal tissue from Barrett's oesophagus through to OAC. Old *et al.* (2017) identified significant biochemical differences categorised by specific spectral signatures using ATR-FTIR. High glycogen content was seen in normal squamous tissue, high glycoprotein content was observed in glandular Barrett's oesophagus tissue, and high DNA content was observed in dysplastic tissue/OAC samples. Classification of normal squamous samples *vs.* 'abnormal' samples (any stage of Barrett's oesophagus) was performed with 100% sensitivity and specificity. Neoplastic Barrett's oesophagus (dysplasia or OAC) was identified with 95.6% sensitivity and 86.4% specificity (Old *et al.*, 2017).

Almond *et al.* (2014) evaluated the ability of endoscopic Raman spectroscopy to objectively detect HGD and OAC. The authors found RS demonstrated a sensitivity of 86% and a specificity of 88% for detecting HGD and adenocarcinoma (*ex vivo* samples $n = 62$ patients).

Bergholt *et al.* (2011) evaluated the biochemical foundation and clinical merit of a multimodal image-guided Raman endoscopy technique for real-time *in vivo* diagnosis of cancer in the oesophagus during OGD. The authors concluded an accuracy of 96.0% (sensitivity 97.0%, specificity 95.2%) for an *in vivo* diagnosis of OAC. These findings may enable multiple points being examined *in vivo* and possibly delineate tumour margins from inflammation.

Raman Spectroscopy has benefits including targeted therapy in endoscopic mucosal resection (EMR) and fast spectral acquisition times (Almond *et al.*, 2011). The technology has shown promise as an adjunct in an *ex vivo* setting aiding pathologists in the absence or presence of dysplasia. Its use in an *in vivo* setting is unclear for dysplasia classification.

Translational studies moving vibrational spectroscopy into the clinical field has been preliminary. Relying on vibrational spectroscopy as a definite diagnostic tool or as a suggestive tool requiring further directive imaging is up for deliberation. Limitations of using vibrational spectroscopy include the need for sensitive and highly optimised instrumentation as well as theoretical heating effects with analysis. A clinician needs to weigh up the advantages with the disadvantages of using the technology prior to advising further direct invasive or non-invasive testing.

1.5 Conclusion

The cost-effectiveness of surveillance in Barrett's oesophagus is often questioned because the rate of conversion from Barrett's oesophagus to adenocarcinoma is only 0.5% per year (Shaheen *et al.*, 2000). Nevertheless, arguments for frequent surveillance include that the majority of adenocarcinoma results within Barrett's oesophagus segments and the risk of OAC is about 30–40 times higher in patients with Barrett's oesophagus compared to those without (Sarr *et al.*, 1985).

Vibrational spectroscopy identifies molecular data that can be rapidly acquired without the need for specialised sample preparation. This could potentially streamline analyses for interventions in many fields in the modern-day NHS. Vibrational spectroscopic techniques have shown promise in delineating progression to malignancy in a number of fields in medicine including gynaecology and neurosurgery.

An earlier diagnosis of dysplasia in Barrett's oesophagus patients would ultimately enable fewer invasive and more expensive surgical options. Optical techniques combined with vibrational spectroscopy could not only aid in differentiating grades of dysplasia in tissue but also identify potential future biomarkers.

1.6 Aims of Thesis

Current literature fails to demonstrate clear evidence for an established benefit and cost-effectiveness of screening and surveillance in Barrett's oesophagus, as well as robust risk stratification for patients to best use NHS resources.

Vibrational spectroscopy can be utilised to identify biomolecular changes that occur prior to any morphological changes. Biochemical and molecular changes evident in biofluids or tissue may enable risk stratification in pre-malignant conditions guiding future surveillance for these patients.

1.6.1 Objectives

- 1) To establish classification models based on ATR-FTIR measurements from plasma, serum, urine, saliva and oesophageal tissue from patients that classify i) normal squamous mucosa from inflammation ii) intestinal metaplasia from LGD iii) LGD from HGD and iv) HGD from OAC
- 2) To establish classification models based on Raman point-based measurements from plasma, serum, urine and saliva and oesophageal tissue from patients that classify i) normal squamous mucosa from inflammation ii) intestinal metaplasia from LGD iii) LGD from HGD and iv) HGD from OAC patients
- 3) To establish spectral differences using Raman spectroscopic mapping between normal squamous epithelia, specialised intestinal metaplasia and OAC in three illustrated cases.

Chapter

2

2. Study demographics

One hundred and twenty-nine (129) patients were recruited into these studies.

The demographics of patients recruited prior to anonymisation between October 2017 and June 2019 are listed in Table 4.

Table 4: Patient demographics

Demographic	Normal (n = 38)	Inflammatory (n = 19)	IM (n = 28)	LGD (n = 6)	HGD (n = 12)	OAC (n = 26)
Median Age (range)	52 (40-88)	56 (38-92)	59 (44-79)	69 (58-81)	69 (55-79)	71 (49-77)
Male sex (n/%)	21 (55)	14 (74)	17 (61)	6 (100)	10 (83)	21 (81)
BMI (n/%)						
- < 24.9	6 (16)	5 (26)	7 (25)	2 (33)	4 (33)	5 (19)
- 25-29.9	31 (82)	12 (63)	21 (75)	3 (50)	8 (67)	20 (77)
- > 30	1 (2)	2 (11)	0 (0)	1 (17)	0 (0)	1 (4)
White Ethnicity	36 (95)	18 (95)	28 (100)	6 (100)	12 (100)	26 (100)
Smoking status						
- Never	16 (42)	6 (32)	17 (60)	3 (50)	5 (42)	7 (27)
- Current	10 (26)	9 (47)	1 (4)	0 (0)	0 (0)	1 (4)
- Ex	12 (32)	4 (21)	10 (36)	3 (50)	7 (58)	18 (69)
History of GORD (n/%)	26 (68)	15 (79)	27 (96)	6 (100)	12 (100)	20 (77)

Patients that tend to be at a higher risk of Barrett's oesophagus and of progression to OAC appear to be of male sex, white ethnicity, increased BMI (>25), ex or current smokers and have a previous or current history of GORD. This correlates with previous epidemiological data (Yachimski *et al.*, 2010; Corley *et al.*, 2009; Corley *et al.*, 2007; Cook *et al.*, 2012).

Further demographics on patient comorbidities were not obtained. To understand the impact of the demographic data a multivariate analysis would need to be performed. This could be considered for future studies.

Chapter

3



Declaration of work

To Whom it May Concern,

Mr Ishaan Maitra helped design the study and collected all biofluids prospectively from targeted patients at Lancashire Teaching Hospitals NHS Foundation Trust. Mr Ishaan Maitra publicised the research, gaining assistance from endoscopy staff and local cancer specialist nurses. Sample processing and storage was kindly undertaken by Katherine L. Ashton.

Mr Ishaan Maitra processed the biofluids using the Bruker TENSOR 27 FTIR spectrometer at the University of Central Lancashire Research laboratory. Data was analysed by Camilo L.M. Morais and Kassio M.G. Lima. Ishaan Maitra, Camilo L.M. Morais and Kassio M.G. Lima put a report together for local endoscopy NHS staff and a paper for publication of the results with the support of Prof. Francis L. Martin.

Signed

.....
Professor F L Martin

.....
Mr I Maitra

Attenuated Total Reflection Fourier Transform Infrared spectral discrimination in oesophageal transformation to adenocarcinoma from human body fluids

Introduction

Multiple methods such as high-resolution endoscopy, narrow band imaging (NBI) and chromoendoscopy have been widely used in the diagnosis and differential diagnosis of oesophageal diseases (Kara *et al.*, 2005 (A); Mannath *et al.*, 2010; Pohl *et al.*, 2010). Disadvantages of current diagnostic methods include the invasive collection of biopsy samples at endoscopy and poor sampling technique despite rigorous protocols. Furthermore, significant inter and intra-observer variability exist for the endoscopist and the histopathologist. The need for simple, reproducible, real time information on an oesophageal disease state by non-invasive methods has never been more relevant.

ATR-FTIR has been used to identify neoplasia in ovarian (Theophilou *et al.*, 2016), cervical (Purandare *et al.*, 2013), breast (Kelly *et al.*, 2011), brain (Hands *et al.*, 2014), prostate (Theophilou *et al.*, 2015), lung (Petibois *et al.*, 2007), skin (Andleed *et al.*, 2018), thyroid (Wu *et al.*, 2016), stomach (Li *et al.*, 2005), colon (Li *et al.*, 2017) and pancreatic tissue (Vazquez-Zapien *et al.*, 2016). Several groups of investigators have used IR spectroscopy for the detection of Barrett's oesophagus and precancerous changes in oesophageal tissues (Maziak *et al.*, 2007; Quaroni and Casson, 2009; Wang *et al.*, 2003; Wang *et al.*, 2007). To the best of our knowledge, ATR-FTIR has never interrogated the oesophageal stages of transformation to OAC from biofluids including plasma, serum, saliva or urine.

Herein we set out to develop an accurate, quick and inexpensive method using biofluids (plasma, saliva, serum or urine) towards detecting oesophageal stages through to OAC (normal; inflammatory; Barrett's oesophagus; low-grade dysplasia (LGD); high-grade dysplasia (HGD); and, OAC). This was using a derived FTIR spectral region, or combination of variables, that reflects a specific biochemical feature of disease in human bodily fluids. We employed successive projections algorithm (SPA) and genetic algorithm (GA) to select an appropriate subset of wavenumbers for quadratic discriminant analysis (QDA).

Materials and Methods

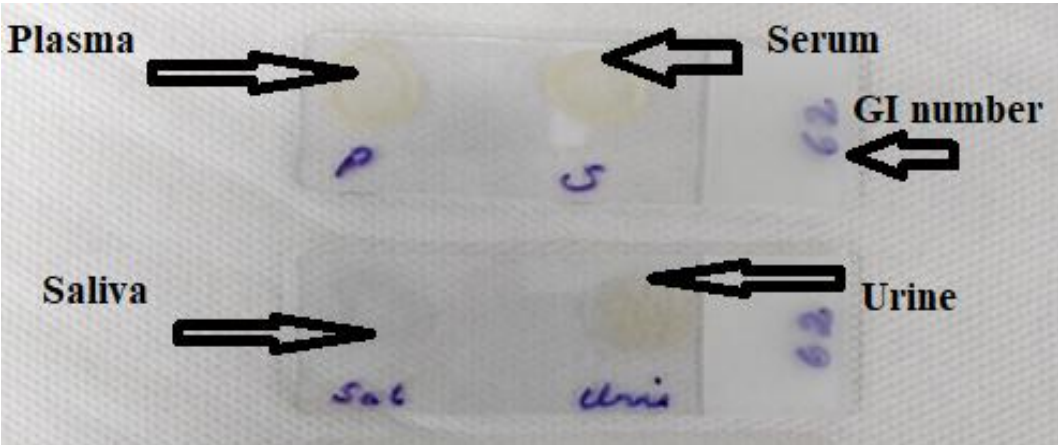
Sample Collection

Patients were identified from upper GI multi-disciplinary team meetings and local hospital pathology databases. Potential participants were identified prospectively and consent for biofluids (blood for plasma or serum, urine or saliva) was taken between October 2017 and June 2019 in a clinic or endoscopy setting. A power test (*t*-test-based with a 95% confidence level) was performed to determine the minimum sample size at 80% power, where a total of 82 samples was suggested. We collected 120 (plasma), 127 (saliva), 124 (serum) and 126 (urine) samples, which surpasses the power of 80%. Biofluid specimens were categorised as follows: i) plasma: $n = 35$ normal, $n = 18$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD and $n = 22$ OAC (set A); ii) saliva: $n = 38$ normal, $n = 19$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD and $n = 25$ OAC (set B); iii) serum: $n = 36$ normal, $n = 19$ inflammatory, $n = 28$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD and $n = 23$ OAC (set C); iv) and, urine: $n = 38$ normal, $n = 19$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 11$ HGD and $n = 25$ OAC (set D).

Ethical approval was granted by the East of England - Cambridge Central Research Ethics Committee from 2015 (Archival gastro-intestinal tissue, blood, saliva and urine collection; REC reference: 18/EE/0069; IRAS project ID: 242639). Ethics was also granted from the University of Central Lancashire (STEMH 909 application). All biofluids were stored in appropriate containers initially in a fridge at 4°C for up to 2h. Prior to freezing, blood samples were centrifuged at 20°C at 2200 rpm for 15 min to obtain plasma and serum samples (local protocol). Saliva samples were taken from patients 3 to 6 h prior to ingestion of solids or liquids. All biofluids were then snap frozen and stored at -80°C to prevent molecular degradation (Baker *et al.*, 2014).

Prior to slide preparation, biofluids samples were left to thaw in the fridge at 4°C to reduce crystal artefact. Urine samples were centrifuged at 20°C at 2200 rpm for 15 min to remove white and red cells prior to pipetting. Thirty μL of individual biofluids (plasma, serum, saliva or urine) were pipetted onto naked FisherBrand™ slides for ATR-FTIR spectral analysis. Each slide was labelled with a specific GI number corresponding to its specific tissue pathological classification (*i.e.*, normal squamous tissue to adenocarcinoma). All slides were left to dry prior to transportation in wooden slide boxes to the laboratory for spectral analysis. All of the samples were stored in a de-humidified glass container to prevent condensation and physical damage.

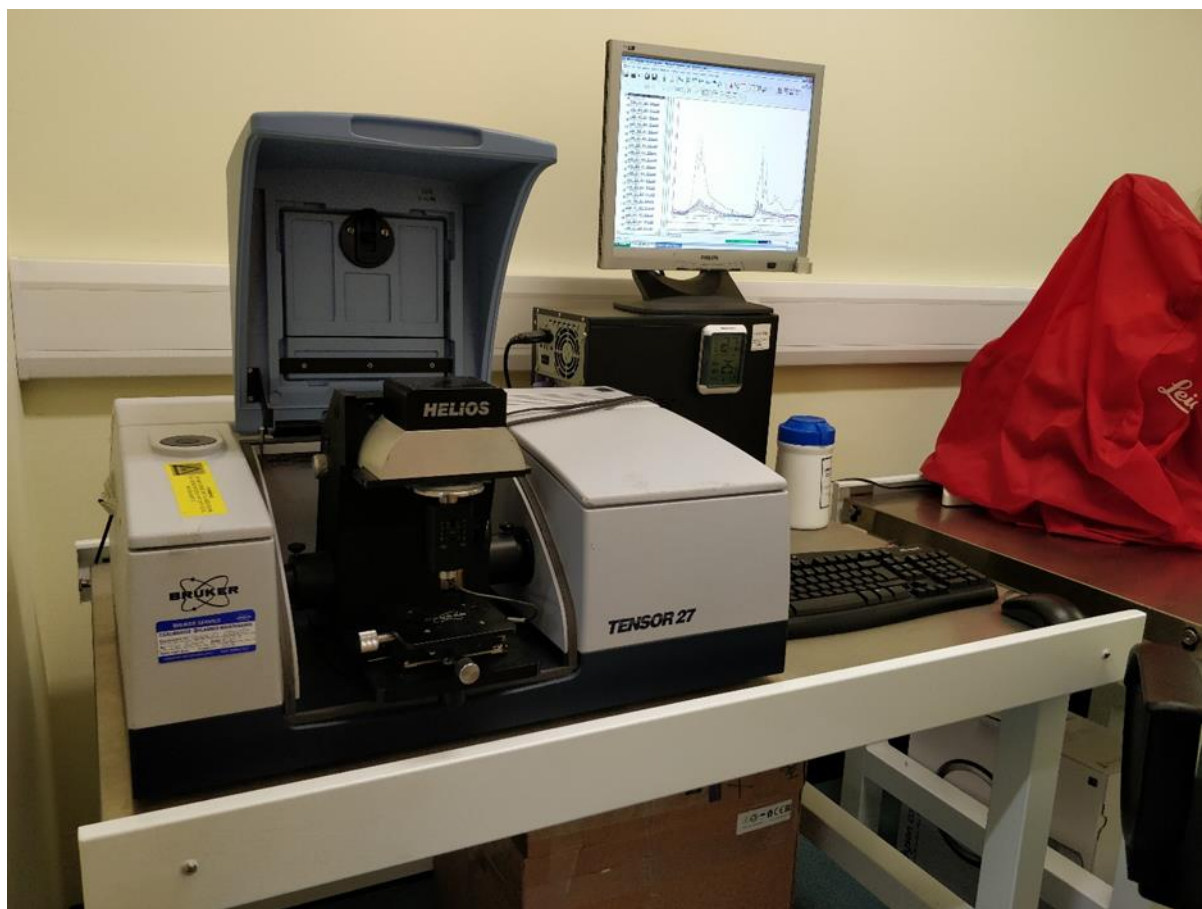
Figure 12: Examples of biofluid slides for ATR-FTIR



ATR-FTIR spectroscopy

Spectroscopic interrogation of biofluid samples was performed at the Biomedical research laboratory at the University of Central Lancashire (UK). Histological diagnoses were unknown to those who performed IR spectroscopy. IR spectra were obtained using a Bruker TENSOR 27 FTIR spectrometer with Helios ATR attachment containing a diamond crystal (Bruker Optics Ltd, Coventry, UK) and operated using OPUS 6.5 software. Spectra were acquired from 10 independent sample locations. Data acquisition parameters were: 8 cm^{-1} spectral resolution giving 4 cm^{-1} data spacing, 32 scans, 6 mm aperture setting and $2\times$ zero-filling factors (Baker *et al.*, 2014). These are standard settings for optimum subcellular interrogation. The ATR diamond crystal was washed with distilled water and dried with tissue paper between each sample and before each new slide. A background absorption spectrum (for atmospheric correction) was taken prior to each new sample (Baker *et al.*, 2014).

Figure 13: Bruker TENSOR 27 FTIR spectrometer



Data analysis and chemometric methods

The data import, pre-treatment and construction of chemometric classification models were implemented in MATLAB R2014a software (MathWorks, USA) by using the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., USA) and custom-made routines. Raw spectra were pre-processed by cutting between 1800 cm^{-1} and 900 cm^{-1} (235 wavenumbers at 4 cm^{-1} spectral resolution), followed by rubberband baseline-correction and normalisation to the Amide I peak (*i.e.*, 1650 cm^{-1}). Before constructing the multivariate classification models (principal component analysis quadratic discriminant analysis, PCA-QDA; successive projections algorithm quadratic discriminant analysis, SPA-QDA; genetic algorithm quadratic discriminant analysis, GA-QDA), the samples were divided into training (60%), validation (20%) and prediction (20%) sets by the classic Kennard–Stone (Kennard and Stone, 1969) uniform sampling algorithm applied to the IR spectra as shown in Table 5. The training samples were used in the modelling procedure, whereas the prediction set was only used in the final classification evaluation. The optimum number of variables for SPA-QDA and GA-QDA was determined according to an average risk G of misclassification. Such a cost function is calculated in the validation set as:

$$G = \frac{1}{N_V} \sum_{n=1}^{N_V} g_n \quad (1)$$

where g_n is defined as:

$$g_n = \frac{r^2(x_n, m_{I(n)})}{\min_{I(m) \neq I(n)} r^2(x_n, m_{I(m)})} \quad (2)$$

and $I(n)$ is the index of the true class for the n th validation object x_n .

In this definition, the numerator is the squared Mahalanobis distance between object x_n (of class index $I(n)$) and the sample mean $m_{I(n)}$ of its true class. The denominator in Eq. (2) corresponds to the squared Mahalanobis distance between object x_n and the centre of the closest incorrect class. The minimum value of the cost function (maximum fitness) is achieved when the selected variables from the original data are as close as possible to its true class and as distant as possible from its incorrect class in the validation set. The GA routine was carried out using 100 generations with 200 chromosomes each. Crossover and mutation probabilities were set to 60% and 1%, respectively. Moreover, the algorithm was repeated three times, starting from different random initial populations. The best solution for GA (in terms of fitness value) resulting from three realisations was employed.

The QDA classification score (Q_{ik}) is estimated using the variance-covariance matrix for each class k and an additional natural logarithm term, as follows:

$$Q_{ik} = (x_i - \bar{x}_k)^T \Sigma_k^{-1} (x_i - \bar{x}_k) + \log_e |\Sigma_k| - 2 \log_e \pi_k \quad (3)$$

where Σ_k is the variance-covariance matrix of class k ; and, $\log_e |\Sigma_k|$ is the natural logarithm of the determinant of the variance-covariance matrix of class k . QDA forms a separated variance model for each class and does not assume that different classes have similar variance-covariance matrices, different to what is assumed by linear discriminant analysis (LDA) (Dixon and Brereton, 2009). The calculation of figures of merit is a recommended standard practice to test model performance (Cheung *et al.*, 2011). Herein, measures of test accuracy including sensitivity (portion of positive samples correctly classified), specificity (portion of negative samples correctly classified), and F-score (a general measurement of the model accuracy) was utilised. These quality metrics are calculated using the following equations:

$$\text{Sensitivity (\%)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (4)$$

$$\text{Specificity (\%)} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \quad (5)$$

$$\text{F-score} = \frac{2 \times \text{SENS} \times \text{SPEC}}{\text{SENS} + \text{SPEC}} \quad (6)$$

where TP stands for true positives, TN for true negatives, FP for false positives and FN for false negatives. SENS stands for sensitivity and SPEC for specificity. All selected wavenumbers derived from SPA-QDA and GA-QDA for oesophageal stages [*i.e.*, normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* low-grade dysplasia (LGD) *vs.* high-grade dysplasia (HGD) *vs.* oesophageal adenocarcinoma (OAC)] were confirmed by a two-tailed Student's *t*-test (95% confidence interval).

Results

The number of training, validation and prediction specimens (or spectra) in each biofluid category is summarised in Table 5.

Table 5: Number of training, validation and prediction specimens (or spectra) in each category of biofluids

Category – biofluids	Training	Validation	Prediction
Normal – plasma	21	7	7
Normal – serum	22	7	7
Normal – saliva	22	8	8
Normal – urine	22	8	8
Inflammatory – plasma	10	4	4
Inflammatory – serum	11	4	4
Inflammatory – saliva	11	4	4
Inflammatory – urine	11	4	4
Barrett’s oesophagus - plasma	17	5	5
Barrett’s oesophagus – serum	18	5	5
Barrett’s oesophagus – saliva	17	5	5
Barrett’s oesophagus – urine	17	5	5
LGD – plasma	3	1	2
LGD – serum	3	1	2
LGD – saliva	3	1	2
LGD – urine	3	1	2
HGD – plasma	7	2	3
HGD – serum	7	2	3
HGD – saliva	7	2	3
HGD – urine	6	2	3
OAC – plasma	12	5	5
OAC – serum	14	4	4
OAC – saliva	15	5	5
OAC – urine	15	5	5

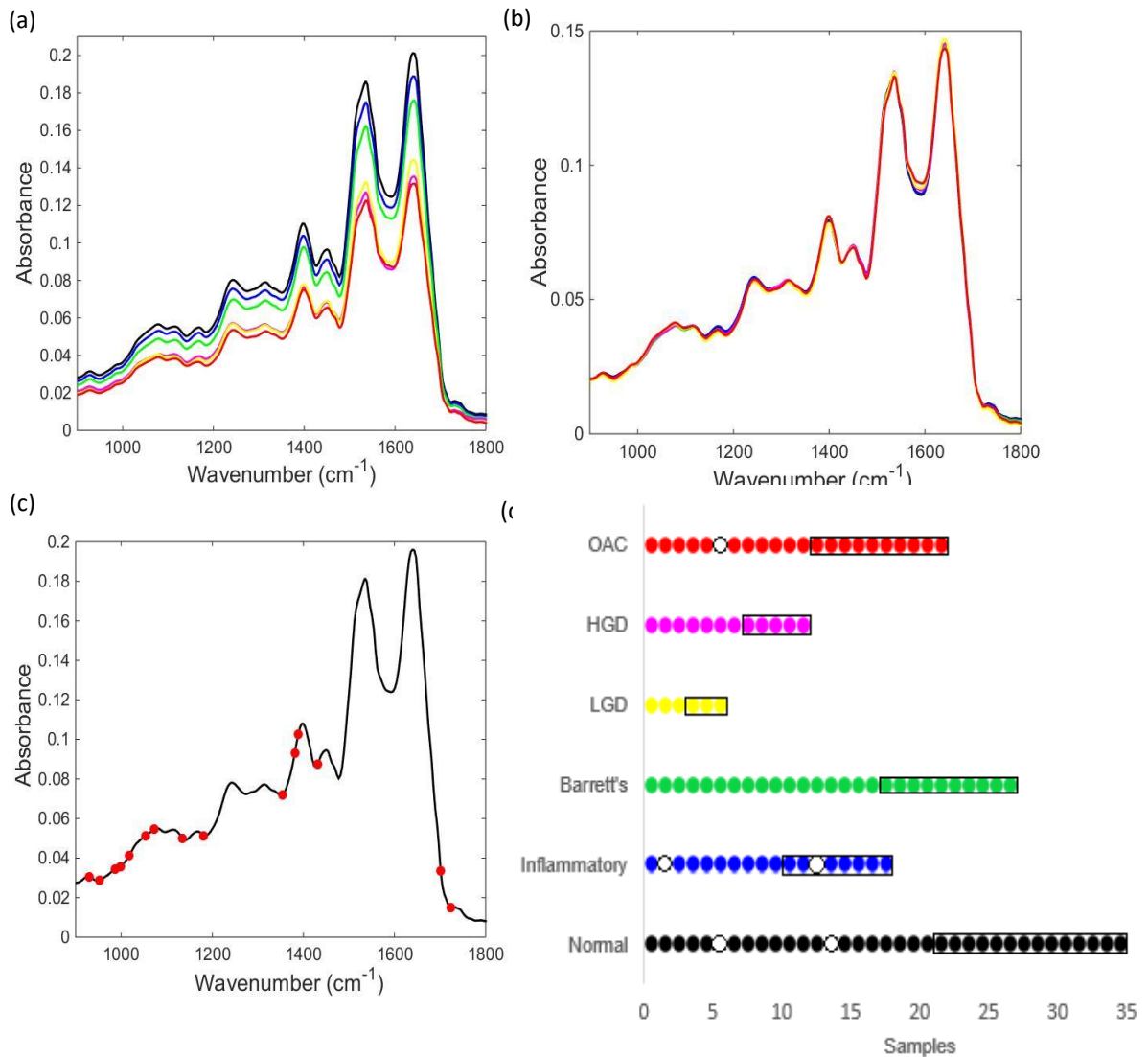
Plasma dataset. Fig. 14A shows the average raw IR spectra derived from blood plasma for the six oesophageal stages of disease (normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC), respectively. Overall, the IR spectra for oesophageal stages appear to overlap in the biochemical-cell fingerprint region (1800 cm^{-1} to 900 cm^{-1}), making it difficult to distinguish any subtle but significant differences. On closer analysis, notable distinguishing peaks that represent protein functional groups could be observed around 1650 cm^{-1} (Amide I) and 1550 cm^{-1} (Amide II). In addition, peaks were observed in the region of 1050 cm^{-1} to 1000 cm^{-1} (carbohydrates and collagen) and 1300 cm^{-1} to 1150 cm^{-1} (Amide III and $\nu_{as}\text{PO}_2^-$). The major peaks of the methyl groups of lipids and proteins could be found at around 1260 cm^{-1} (Amide III), 1225 cm^{-1} ($\nu_{as}\text{PO}_2^-$) and 1080 cm^{-1} ($\nu_s\text{PO}_2^-$). To discriminate the six oesophageal stages, the spectral dataset was pre-processed using baseline correction and normalisation using the Amide I peak (Figure 14B). Average IR pre-processed spectra again appear to overlap in the biochemical fingerprint region (1800 cm^{-1} to 900 cm^{-1}) making spectral observation and the discovery of markers or signatures difficult. Therefore, chemometric techniques such as PCA-QDA, SPA-QDA and GA-QDA algorithms were adopted to classify normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC based on their IR spectra.

A classification of the six oesophageal stages was developed by discriminant analysis using the IR spectra between 900 cm^{-1} and 1800 cm^{-1} . Figure 14C shows the wavenumbers associated with class differences. The GA-QDA model utilised in plasma samples was able to predict all stages of disease to OAC with 100% accuracy, sensitivity and specificity. Fig. 14D shows the predicted class achieved by the GA-QDA model for all classes based on 15 selected wavenumbers. An excellent classification of the samples was achieved (only 3 errors in the training set and 1 error in the prediction set). The PCA-QDA model using the scores on seven PCs (90% of the total data variance) achieved 100% for accuracy, sensitivity, specificity and F-scores for normal, inflammatory, Barrett's oesophagus and OAC classes as can be seen in Table 6. SPA-QDA also achieved a very high accuracy for classification for LGD (100%) using 7 wavenumbers (namely 1392 cm^{-1} , 1485 cm^{-1} , 1539 cm^{-1} , 1585 cm^{-1} , 1624 cm^{-1} , 1643 cm^{-1} and 1681 cm^{-1}), as shown in Table 6. Table ‡S1 (Appendix) lists the selected wavenumbers obtained for SPA-QDA and GA-QDA models for plasma samples with their tentative biomolecular assignments. Using a student's t-test (95% confidence interval), key discriminant wavenumbers in plasma categorising disease processes to OAC were between 929 cm^{-1} to 1431 cm^{-1} , associated with DNA/RNA and proteins.

Table 6: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA and SPA-QDA) results for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using plasma samples.

FOM	PCA – QDA model for oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	100	100	100	92.3	92.3	100
Sensitivity (%)	100	100	100	100	91.3	100
Specificity (%)	100	100	100	0	100	100
F-Scores (%)	100	100	100	0	95.4	100
FOM	SPA – QDA model for Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	84.6	84.6	96.1	100	96.7	92.3
Sensitivity (%)	94.7	86.3	95.2	100	100	95.2
Specificity (%)	57.1	75.0	100	100	66.6	80
F-Scores (%)	71.2	80.2	97.5	100	80	86.9

Figure 14: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using plasma samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 15 selected wavenumbers by GA-QDA. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ – illustrates a misclassification.



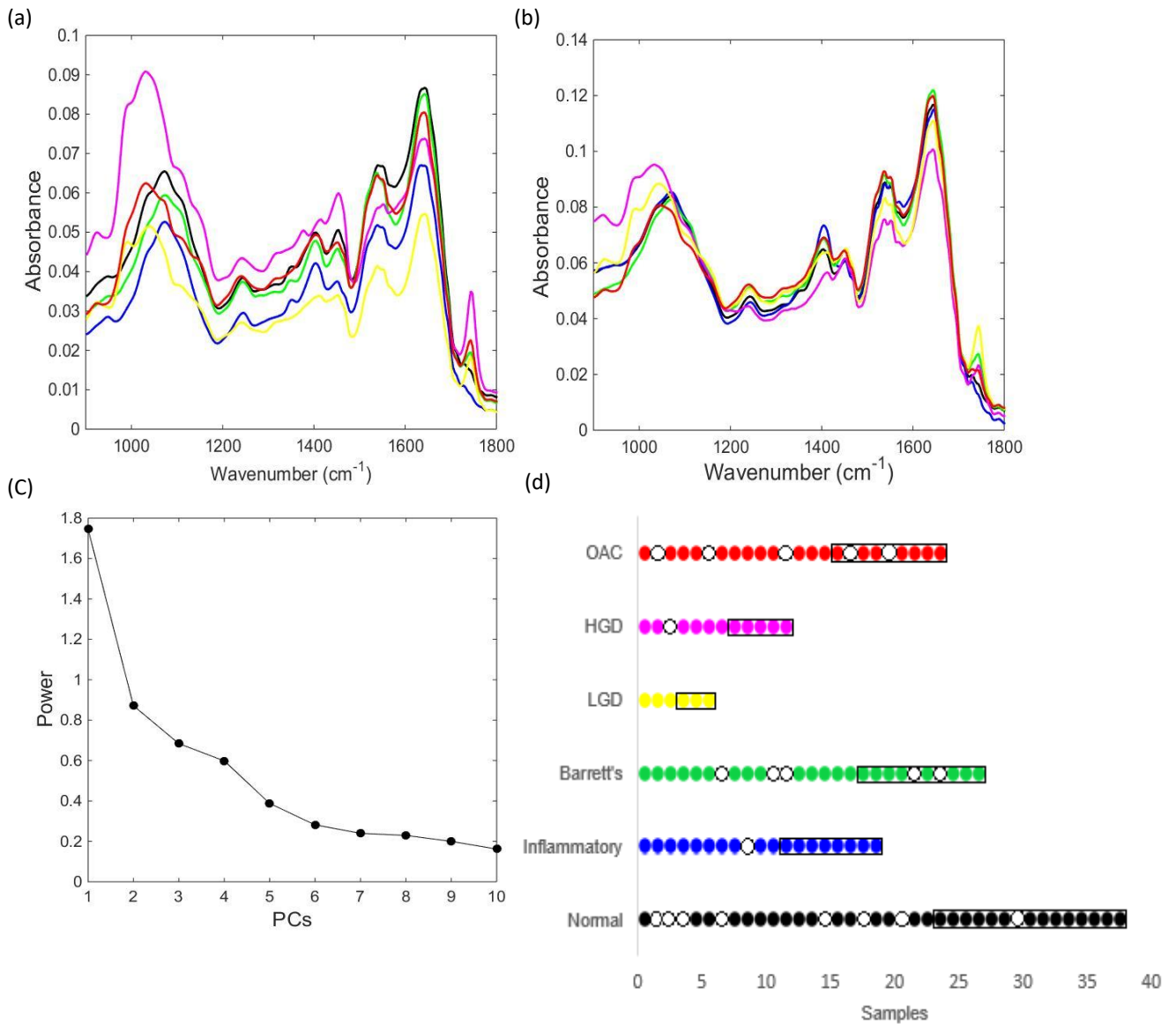
Saliva dataset. Figure 15A and 15B shows the raw average IR spectra and average pre-processed (baseline correction and normalisation) in the biochemical-cell fingerprint region (1800 cm^{-1} to 900 cm^{-1}) derived from saliva for the six oesophageal stages. There are notable differences in the wavenumber regions 1000 cm^{-1} to 1150 cm^{-1} (DNA/RNA region), 1350 cm^{-1} to 1500 cm^{-1} (Amide II) and 1530 cm^{-1} to 1600 cm^{-1} (Amide I) for all classes investigated. Although there is some overlap amongst the average pre-processed spectra, the application of chemometric techniques (PCA-QDA, SPA-QDA or GA-QDA) exhibits good categorisation for all the classes (normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC).

As shown in Table 7, classification of the six oesophageal stages using saliva and the FTIR spectra between 900 cm^{-1} and 1800 cm^{-1} using all models (PCA-QDA, SPA-QDA and GA-QDA) gives excellent classification accuracy (88.8–100%). There was poor specificity in the LGD class using PCA-QDA and GA-QDA. SPA-QDA is the best classification model for saliva, with F-Score values ranging from 76.9 to 100%. PCA-QDA achieved 100% for all figures of merit for the inflammatory category in the prediction set. The accuracy and sensitivity were found to be $>88\%$ for all the other classes, using the scores on seven PCs from PCA (Figure 15C). The number of PCs is selected in Figure 15C based on the minimum number of PCs that generates the lowest power (eigenvalue), before the power follows a constant trend. Fig. 15D shows the predicted class achieved for the PCA-QDA model for all classes. A good classification of the samples was observed (15 errors in the training set and 5 errors in the prediction set). For SPA-QDA, 7 selected wavenumbers (902 cm^{-1} , 1014 cm^{-1} , 1099 cm^{-1} , 1589 cm^{-1} , 1643 cm^{-1} , 1697 cm^{-1} , and 1743 cm^{-1}) provided excellent classification, especially for the LGD class (100% for all FOMs). The classification performance using GA-QDA was 100% for all figures of merit for the Barrett's oesophagus class based on 14 selected wavenumbers (991 cm^{-1} , 1003 cm^{-1} , 1068 cm^{-1} , 1107 cm^{-1} , 1431 cm^{-1} , 1558 cm^{-1} , 1585 cm^{-1} , 1604 cm^{-1} , 1624 cm^{-1} , 1689 cm^{-1} , 1701 cm^{-1} , 1716 cm^{-1} , 1778 cm^{-1} and 1786 cm^{-1}). Table S2 (see SI) lists the selected wavenumbers obtained by SPA-QDA and GA-QDA models for saliva samples with their tentative biomolecular assignments.

Table 7: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using saliva samples.

FOM – PCA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	92.6	100	96.3	88.8	92.6	92.6
Sensitivity (%)	89.4	100	100	96	91.6	100
Specificity (%)	100	100	80	0	100	75
F-Scores (%)	94.4	100	88.9	0	95.6	75
FOM – SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	88.8	85.2	96.2	100	96.3	96.3
Sensitivity (%)	100	86.9	95.4	100	100	95.4
Specificity (%)	62.5	75	100	100	66.6	100
F-Scores (%)	76.9	80.5	97.6	100	80	97.7
FOM – GA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	92.5	92.6	100	92.6	88.9	96.2
Sensitivity (%)	100	91.3	100	100	91.6	95.4
Specificity (%)	75	100	100	0	66.6	100
F-Scores (%)	85.7	95.4	100	0	77.1	97.7

Figure 15: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using saliva samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) Singular value decomposition (SVD) vs. Number of principal component (PC) obtained by PCA-QDA, where Power represents the eigenvalue. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for PCA-QDA. \circ - illustrates a misclassification.

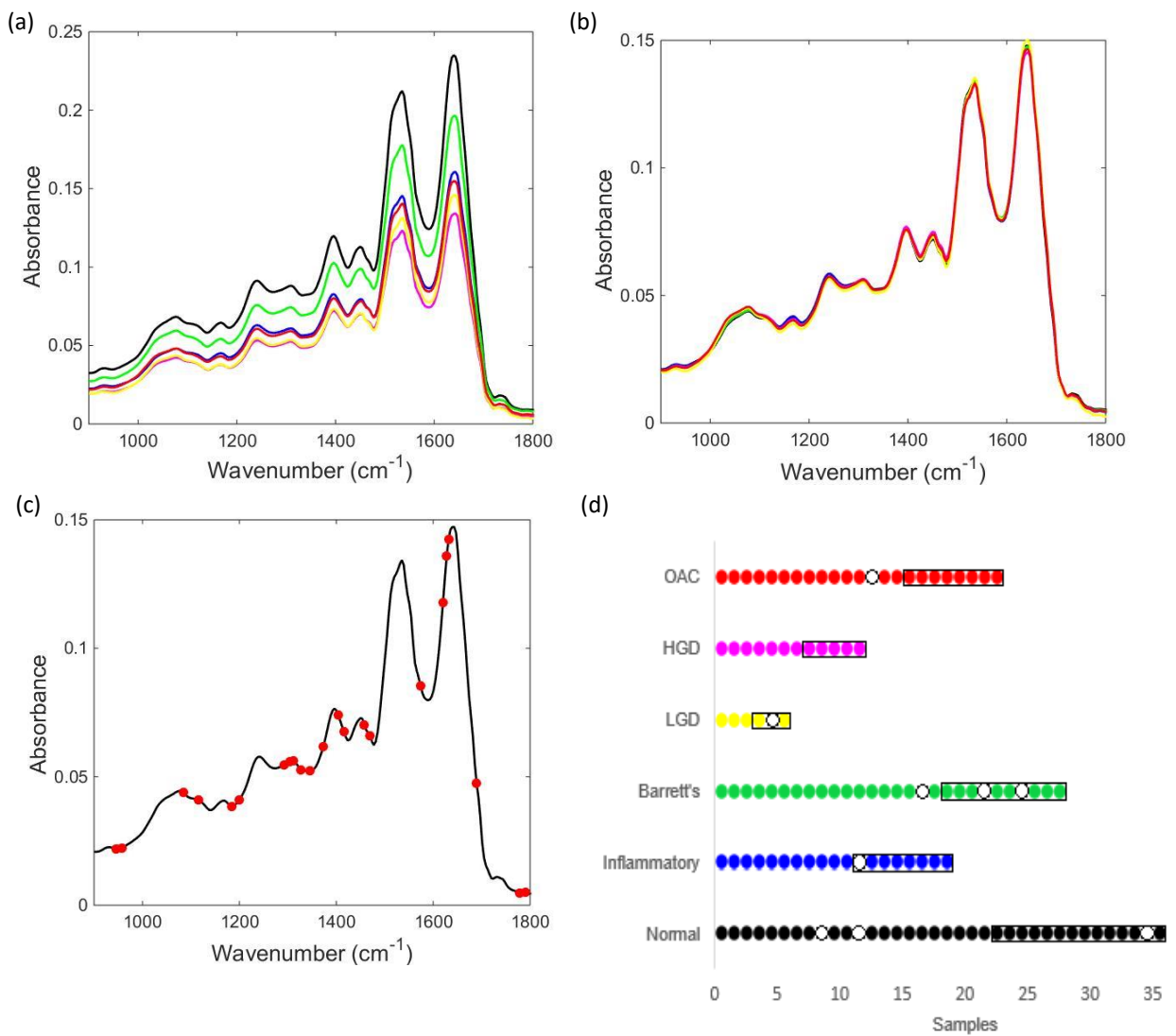


Serum dataset. Figures 16A and 16B shows the average raw and pre-processed spectra for serum samples obtained from the six oesophageal stages. Classification of normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC for serum was performed using discriminant analysis of the spectral bio-fingerprint region (1800 cm^{-1} - 900 cm^{-1}). GA-QDA, using only 13 selected wavenumbers (1000 cm^{-1} , 1315 cm^{-1} , 1319 cm^{-1} , 1330 cm^{-1} , 1338 cm^{-1} , 1435 cm^{-1} , 1442 cm^{-1} , 1446 cm^{-1} , 1492 cm^{-1} , 1539 cm^{-1} , 1573 cm^{-1} , 1600 cm^{-1} and 1654 cm^{-1} ; Fig. 16C) achieved 100% for all figures of merit for the OAC stage, as shown in Table 8. The GA-QDA model demonstrated an excellent classification with 4 errors in the training set and 4 errors in the prediction set (Figure 16D). For normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* OAC stages, the SPA-QDA model demonstrated sensitivities and specificities more than 71% (Table 8) using only 7 wavenumbers (1041 cm^{-1} , 1477 cm^{-1} , 1539 cm^{-1} , 1593 cm^{-1} , 1631 cm^{-1} , 1662 cm^{-1} and 1743 cm^{-1}). The PCA-QDA models for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* OAC stages produced sensitivity and specificity greater than 71% using seven PC scores, which accounted for more than 90% of original data variance. Some notable differences using a student's t-test can be observed in the wavenumber regions between 1000 cm^{-1} to 1338 cm^{-1} (adenine vibration of DNA/RNA region), 1435 cm^{-1} to 1573 cm^{-1} (methyl groups of proteins and amide II absorption) and 1600 cm^{-1} to 1654 cm^{-1} (Amide I) for all of the classes investigated. Table ‡S3 (see SI) lists the selected wavenumbers obtained by SPA-QDA and GA-QDA models for serum samples with their tentative biomolecular assignments.

Table 8: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using serum samples.

FOM – PCA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	92.0	92.0	96.0	92.0	92.0	96.0
Sensitivity (%)	100	90.4	100	95.6	95.4	95.2
Specificity (%)	71.4	100	80	50.0	66.6	100
F-Scores (%)	83.3	95.0	88.8	65.6	78.5	97.5
FOM – SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	92.0	88.0	92.0	92.0	92.0	96.0
Sensitivity (%)	100	90.4	95.0	95.6	95.4	95.2
Specificity (%)	71.4	75.0	80.0	50.0	66.6	100
F-Scores (%)	93.3	82.0	86.8	65.6	78.5	97.5
FOM – GA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	96.0	96.0	96.0	92.0	96.0	100
Sensitivity (%)	100	95.2	100	95.6	95.4	100
Specificity (%)	85.7	100	80.0	50.0	100	100
F-Scores (%)	92.3	97.5	88.8	65.6	97.6	100

Figure 16: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using serum samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 13 selected wavenumbers by GA-QDA. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.



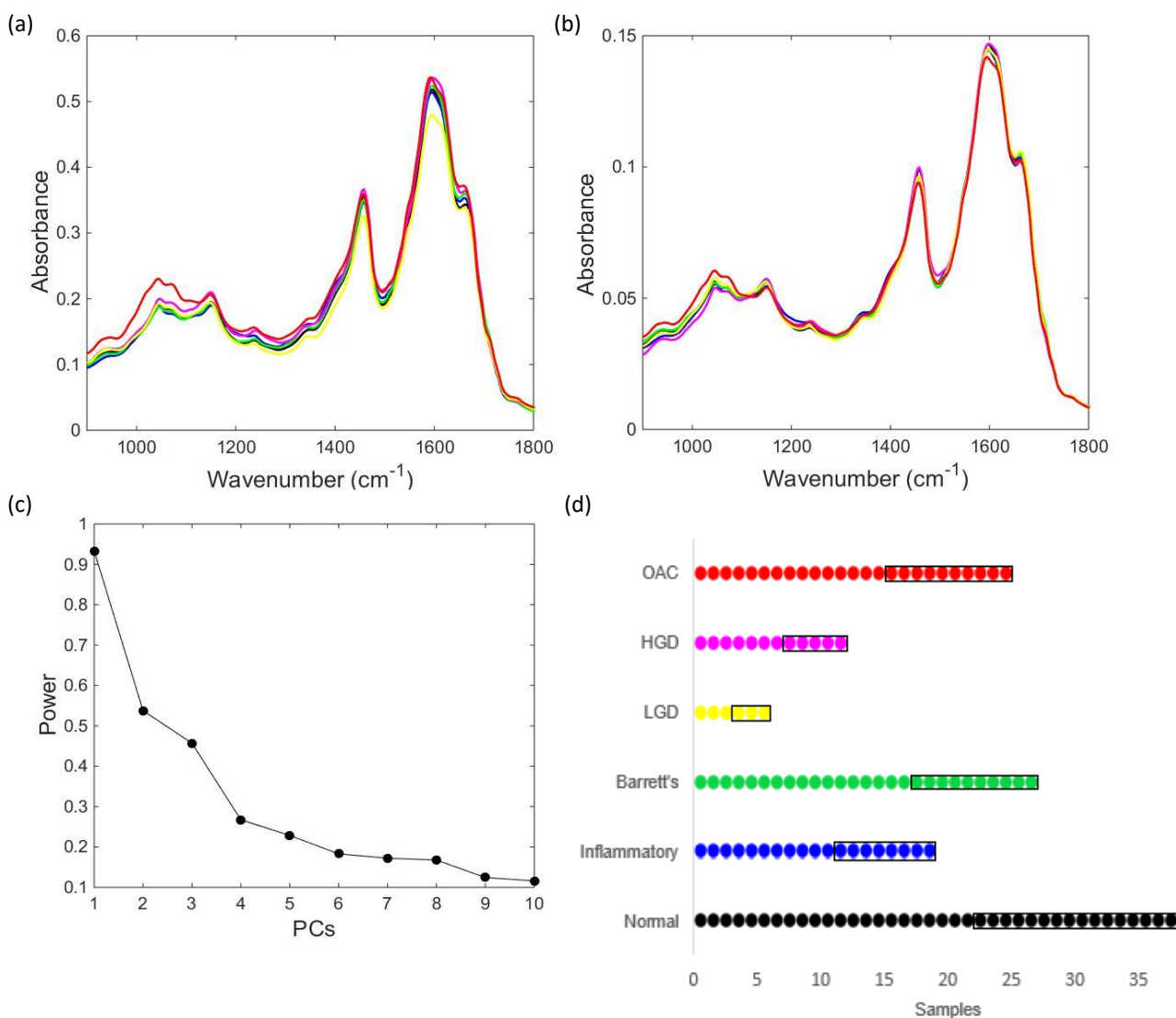
Urine dataset. Figure 17A and 17B shows the raw and average pre-processed (baseline correction and normalisation) spectra within the bio-fingerprint region (1800 cm^{-1} to 900 cm^{-1}) derived from urine for all the six oesophageal stages. Although the discrimination for all classes on the basis of IR spectra was difficult due to the complexity of the spectra, good models using PCA-QDA, SPA-QDA and GA-QDA were found for urine. The PCA-QDA model using urine spectra achieved 100% of classification for all figures of merit (accuracy, sensitivity, specificity and F-scores) for test samples using seven PC scores as shown in Figure 17C. Fig. 17D demonstrated that the predicted classification performance achieved by the PCA-QDA model for all classes was 100% (no errors in training and prediction sets). As demonstrated in Table 9, both SPA-QDA and GA-QDA models for urine samples achieved high sensitivity values for all classes (78.2–100%). Poor Specificities were poor for normal squamous epithelial samples (35.5%), HGD (33.3%) and OAC (20.0%) using the SPA-QDA model.

The SPA-QDA model using only 7 selected wavenumbers (956 cm^{-1} , 1242 cm^{-1} , 1577 cm^{-1} , 1600 cm^{-1} , 1651 cm^{-1} , 1681 cm^{-1} and 1712 cm^{-1}) achieved good classification, especially for inflammatory, Barrett's oesophagus and LGD classes (F scores $> 76.6\%$). The predicted classification rate using GA-QDA demonstrates good results ($> 80\%$ for all FOM) for Barrett's oesophagus and HGD classes based on 19 selected wavenumbers (956 cm^{-1} , 995 cm^{-1} , 1018 cm^{-1} , 1030 cm^{-1} , 1095 cm^{-1} , 1118 cm^{-1} , 1141 cm^{-1} , 1253 cm^{-1} , 1334 cm^{-1} , 1381 cm^{-1} , 1431 cm^{-1} , 1446 cm^{-1} , 1500 cm^{-1} , 1550 cm^{-1} , 1562 cm^{-1} , 1681 cm^{-1} , 1724 cm^{-1} , 1735 cm^{-1} and 1777 cm^{-1}). Discriminant spectral differences can be observed in the wavenumber regions between 956 cm^{-1} to 1381 cm^{-1} (DNA/RNA region), 1431 cm^{-1} to 1562 cm^{-1} (Amide II stemming from C-N stretching and C-N-H bending vibrations) and 1681 cm^{-1} to 1777 cm^{-1} (a high frequency vibration of an antiparallel β -sheet of Amide I and stretching vibration of lipids). Table ‡S4 (see SI) lists the selected wavenumbers obtained by SPA-QDA and GA-QDA models for urine samples with their tentative biomolecular assignments.

Table 9: ROC analysis (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using urine samples.

FOM – SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	66.6	77.7	92.6	96.2	92.6	77.7
Sensitivity (%)	78.9	78.2	95.4	96.0	100	90.9
Specificity (%)	35.5	75.0	80.0	100	33.3	20.0
F-Scores (%)	50.8	76.6	87.1	97.9	50.0	32.7
FOM – GA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	74.1	85.2	92.6	92.6	96.3	85.2
Sensitivity (%)	78.9	86.9	95.4	96.0	95.8	100
Specificity (%)	62.5	75.0	80.0	50.0	100	20.0
F-Scores (%)	69.7	80.5	87.0	65.7	97.8	33.3

Figure 17: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using urine samples. The panel shows: (A) Average raw spectra in mid-IR region of 1800 cm^{-1} to 900 cm^{-1} and (B) average pre-processed IR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) Singular value decomposition (SVD) vs. Number of principal component (PC) obtained by PCA-QDA, where Power represents the eigenvalue. (D) Predicted class plot vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for PCA-QDA. \circ - illustrates a misclassification.



Discussion

There is no standard surveillance program for the early detection of adenocarcinoma in the oesophagus. The most reliable diagnostic measurement uses endoscopy with biopsy, which is advocated for screening oesophageal neoplasia especially in known high-risk patients. Although endoscopy has been widely used in the diagnosis of oesophageal pathology, approximately 20% of early oesophageal cancer lesions are barely visible to the naked eye (Shah and Gerdes, 2015). In order to have a robust surveillance program for early detection and improved cure rate for OAC, a skilled endoscopist and pathologist are necessary for proper mucosal sampling and histopathologic examination. The development of a quick, convenient, and inexpensive method for detecting early cancer or different stages of specialised intestinal metaplasia to OAC can be useful specifically to guide tissue biopsy, thus increasing the yield of dysplasia detection.

An approach to OAC screening in the general population based on biofluids (blood plasma, serum, saliva and urine) interrogated by ATR-FTIR spectroscopy linked with feature selection methods for classification could be the potential to segregate stages to oesophageal adenocarcinoma. The use of ATR-FTIR spectroscopy coupled multivariate classification techniques (PCA-QDA, SPA-QDA and GA-QDA) in identifying oesophageal stages of disease to adenocarcinoma has achieved excellent accuracy, sensitivity and specificity, encouraging investigation of screening for others cancers with known markers.

ATR-FTIR spectroscopy was employed to predict six oesophageal stages in four different biofluids (plasma, saliva, serum and urine). PCA-QDA and GA-QDA models were found to give the best class differentiation compared to the SPA-QDA. The GA-QDA model utilised in plasma samples was able to predict all stages of disease to OAC with 100% accuracy, sensitivity and specificity. For this model, several selected wavenumbers appear to be of particular interest, especially at 999 cm^{-1} and 1381 cm^{-1} , representing the ring stretching vibrations mixed strongly with CH in-plane bending and C-O stretching, respectively. Other selected wavenumbers found by GA-QDA using plasma can be found in Table ‡S1. In general, the spectral alterations responsible for the discrimination of oesophageal stages based on plasma were mainly associated with DNA/RNA and proteins at wavenumbers between 929 cm^{-1} and 1431 cm^{-1} . The major advantage for this model, using blood plasma, is the minimal sample preparation for FTIR analysis. The selected wavenumbers obtained for GA-LDA for all classes are related with biochemical signatures the disease besides high accuracy for classification rates.

Since a direct, easy-to-use test to diagnose oesophageal stages to OAC in human saliva is currently not available, we suggest ATR-FTIR spectroscopy combined with multivariate classification techniques for the development of a direct test that meets these challenges. Although saliva has complex biology (containing water, proteins and inorganic substances), we believe that saliva analysis based on ATR-FTIR is a powerful diagnostic tool in OAC. The PCA-QDA model demonstrated good figures of merit ($> 75\%$) in diagnosing all disease states to OAC except LGD. This would potentially limit its use in clinical practice. The selected wavenumbers found by SPA-QDA and GA-QDA models using saliva samples can be found in Table 9 (see SI). Several selected wavenumbers appear to be of particular interest in saliva, namely, the wavenumbers at 1604 cm^{-1} , 1624 cm^{-1} and 1643 cm^{-1} , representing the adenine vibration in DNA, peak of nucleic acids due to the base carbonyl stretching, and ring breathing mode and amide I band (arising from C=O stretching vibrations), respectively.

The classification rates for oesophageal disease using chemometric analysis for plasma were generally found to be better than those of saliva and serum for all classes. One explanation for such a difference is that plasma contains thousands of biomolecules at various concentrations while serum and saliva may contain a more limited number of analytes such as antibodies, proteins and electrolytes resulting in fewer variations in peak intensities or shifts of the IR spectrum.

Problems can arise with liquid samples, such as urine, because very low concentration components may not be detected. Low concentrations can be overcome by drying samples directly on to the crystal to increase their concentrations and, therefore, increase their signal intensities. We believe that clinically relevant levels could be detected in ATR-FTIR spectra of dried insoluble fractions of urine samples without any requirement for chemical manipulation. Urine can be collected non-invasively and without the need for a trained professional to be present. It can also be collected frequently and stored for several days. It is therefore ideal as a diagnostic medium if it can provide clinically useful information. Urine samples for this study demonstrated excellent discrimination for all stages using urine samples based on PCA-QDA models (100% for all FOM). SPA-QDA and GA-QDA presented satisfactory segregation amongst all classes using a few selected wavenumbers, as can be seen in Table 9. The SPA-QDA model produced low specificity values ($<50\%$) for normal, HGD, and OAC disease categories (Table 9).

The classification performance of the spectroscopy-based models to determine oesophageal stages are a result of a combination between the algorithm being used and the nature of the sample being measured. Although the type of classifier employed in this study is the same for all algorithms (QDA), the feature extraction (PCA) and feature selection (SPA and GA) methods work in a different manner, thus leading the models to different results. PCA reduces the spectral dataset to features representing the main sources of variance in the data, but these sources of variation are not necessarily correlated with differences between the samples. The SPA and GA algorithms work by reducing data collinearity (SPA – reduces strong association between two potential predictor variables) and by mimicking the process of natural selection in a computational fashion (GA) (Morais *et al.*, 2019). Both SPA and GA act on the original sample space, whereas PCA projects the sample on an orthogonal space. Choosing the right algorithm for data analysis is an empirical process and is key when interpreting results.

Eight biofluid samples (plasma n = 2; serum n = 2; saliva n = 2; urine n = 2) used for analysis from patients classified in OAC group had chemotherapy prior to biofluid collection. Currently, there are no validated studies suggesting that there are significant IR spectral differences associated with chemotherapy in biofluids (Sala *et al.*, 2020). The addition of these samples in the analysis would not statistically affect the average raw class spectral data for OAC for each biofluid and thus be insignificant with regards to the figures of merit created by the predictive chemometric models used.

Molecular signatures of oesophageal adenocarcinoma and Barrett's oesophagus have been found in urine (Davis *et al.*, 2012); and other biofluids, such as blood (Kunzmann *et al.*, 2018) and saliva (Rapado-González *et al.*, 2016). For this reason, and confirmed by the results reported herein, we believe that IR spectroscopy can be used to discriminate oesophageal transformation to adenocarcinoma based on these biofluids.

Conclusion

The results of this study show that ATR-FTIR spectroscopy coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA), result in a powerful alternative approach for the detection of oesophageal stages of disease to OAC. Herein, we present a new, rational and convenient approach to different biofluids (plasma, saliva, serum and urine) using ATR-FTIR spectroscopy, opening a new level of non-invasive diagnostic tool in this field. We demonstrate a fast, clean, and non-destructive methodology involving minimal sample preparation to categorise the samples. The GA-QDA model utilised in plasma samples was able to predict all stages of disease to OAC with 100% accuracy, sensitivity and specificity. For urine samples, the resulting PCA-QDA model successfully detects biochemical alterations at the maximum classification rate (100%) for different figures of merit (accuracy, sensitivity, and F-scores) for all disease states to OAC. These method in these biofluids make it possible to detect all the oesophageal stages to adenocarcinoma without special sample preparation and reagents, from a minimal sample volume and (almost) immediately after sample collection. In this pilot study, we have demonstrated for the first time that saliva- and urine-based ATR-FTIR spectroscopy coupled with a multivariate classification algorithm has the potential to discriminate oesophageal stages. Further work with biofluids and spectroscopic analysis should be performed in the future to validate these results.

Chapter

4

uclan



Declaration of work

To Whom it May Concern,

Mr Ishaan Maitra helped design the study and collected all biofluids prospectively from targeted patients at Lancashire Teaching Hospitals NHS Foundation Trust. Mr Ishaan Maitra publicised the research, gaining assistance from endoscopy staff and local cancer specialist nurses. Sample processing and storage was kindly undertaken by Katherine L. Ashton.

Mr Ishaan Maitra processed the biofluids using the Renishaw 1000 system Raman spectrometer at the University of Central Lancashire Research laboratory. Data was analysed by Camilo L.M. Morais and Kassio M.G. Lima. Ishaan Maitra, Camilo L.M. Morais and Kassio M.G. Lima prepared a paper for publication of the results with the support of Prof. Francis L. Martin.

Signed

.....

Professor F L Martin

.....

Mr I Maitra

Raman spectral discrimination in human liquid biopsies of oesophageal transformation to adenocarcinoma

Introduction

Raman spectroscopy is complementary to infrared (IR) technology and has advantageous properties when analysing biofluids. Contrary to conventional detection methods, optical spectroscopy on plasma, serum, saliva or urine provides an opportunity to diagnose diseases non-invasively. Raman spectroscopy is based on inelastic vibrational scattering, which can detect the secondary constitution of molecules (Butler *et al.*, 2016). No labelling is necessary and the technique provides high spectral sensitivity (Parker, 1983). Spectroscopy on biofluids has been focused on plasma and serum analysis due to the large readily available biobanks in research laboratories. Multiple studies have been performed demonstrating the potential of Raman spectroscopy for differentiating normal subjects from patients with colorectal (Li *et al.*, 2012), hepatocellular (Taleb *et al.*, 2013), cervical (González-Solís *et al.*, 2014), and breast cancers (Pichardo-Molina *et al.*, 2007) from plasma and serum. Raman spectroscopy in serum has been performed to differentiate controls from oral cancers (Sahu *et al.*, 2013); however, to date, no research has been established using Raman spectroscopy to investigate oesophageal transformation to adenocarcinoma from human body fluids.

This chapter proposes an accurate, fast, and inexpensive method using biofluids (plasma, saliva, serum and urine) for detecting oesophageal stages through to OAC (normal; inflammatory; Barrett's oesophagus; low-grade dysplasia (LGD), high-grade dysplasia (HGD); and, OAC) using Raman spectroscopy.

Materials and Methods

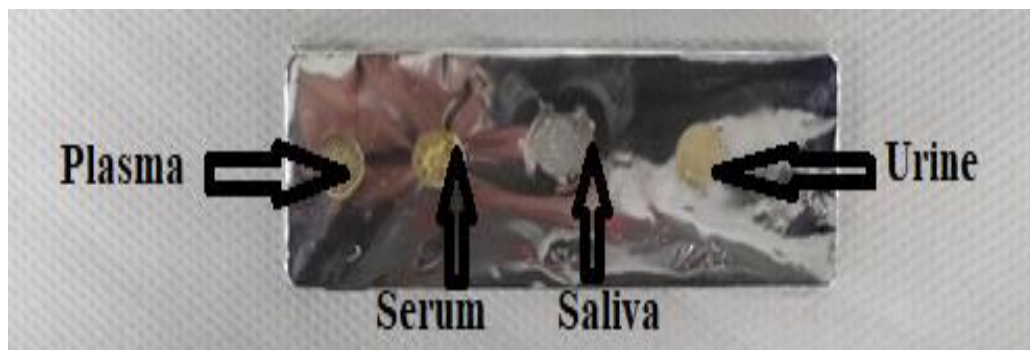
Sample collection

Patients were identified from upper GI multi-disciplinary team meetings and pathology hospital databases which had been created by the pathology laboratory manager and one of the Consultant GI histopathologists. Potential patients were identified prospectively and consent for biofluids (blood for plasma and serum; urine and saliva) was taken between October 2017 and June 2019 in a clinic or endoscopy setting. A power test (t-test-based with a 95% confidence level) was performed to determine the minimum sample size at 80% power, where a total of 82 samples was suggested. We collected 120 (plasma), 118 (saliva), 124 (serum) and 127 (urine) samples, which surpasses the power of 80%. The biofluid specimens were categorised as follows: i) plasma: $n = 35$ normal, $n = 18$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD and $n = 22$ OAC (set A); ii) saliva: $n = 35$ normal, $n = 18$ inflammatory, $n = 26$ Barrett's oesophagus, $n = 5$ LGD, $n = 10$ HGD and $n = 24$ OAC (set B); serum: $n = 36$ normal, $n = 19$ inflammatory, $n = 28$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD and $n = 23$ OAC (set C); and, urine: $n = 38$ normal, $n = 19$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 11$ HGD and $n = 26$ OAC (set D).

Ethical approval was granted by the East of England - Cambridge Central Research Ethics Committee from 2015 (Archival gastro-intestinal tissue, blood, saliva and urine collection; REC reference: 18/EE/0069; IRAS project ID: 242639). Ethics was also granted from the University of Central Lancashire (STEMH 909 application). All biofluids taken at source patient contact were stored in their containers in the fridge at 4°C to 7°C. Prior to freezing, blood samples were centrifuged at 20°C at 2200 rpm for 15 min to obtain plasma and serum samples (local protocol). Saliva samples were taken from patients 3 to 6 h prior to ingestion of solids or liquids. All biofluids were then snap frozen and stored at -80°C to prevent molecular degradation.

Prior to slide preparation, biofluids samples were left to thaw in the fridge at 4°C to prevent crystal artefact. Urine samples were centrifuged at 20°C at 2200 rpm for 15 min to remove white and red cells prior to pipetting. Thirty mL of individual biofluids (plasma, serum and saliva) were pipetted onto aluminium foil-lined FisherBrand™ slides for Raman spectroscopy analysis. This is standard biofluid preparation protocol prior to biofluid Raman spectroscopy analysis (Butler *et al.*, 2016). Each slide was labelled with a specific GI number used to anonymise samples. All slides were left to dry prior to transportation in wooden slide boxes to the spectroscopy laboratory for analysis. Samples were stored in a de-humidified glass container to prevent condensation and physical damage.

Figure 18: Example of biofluid slide for Raman spectroscopy



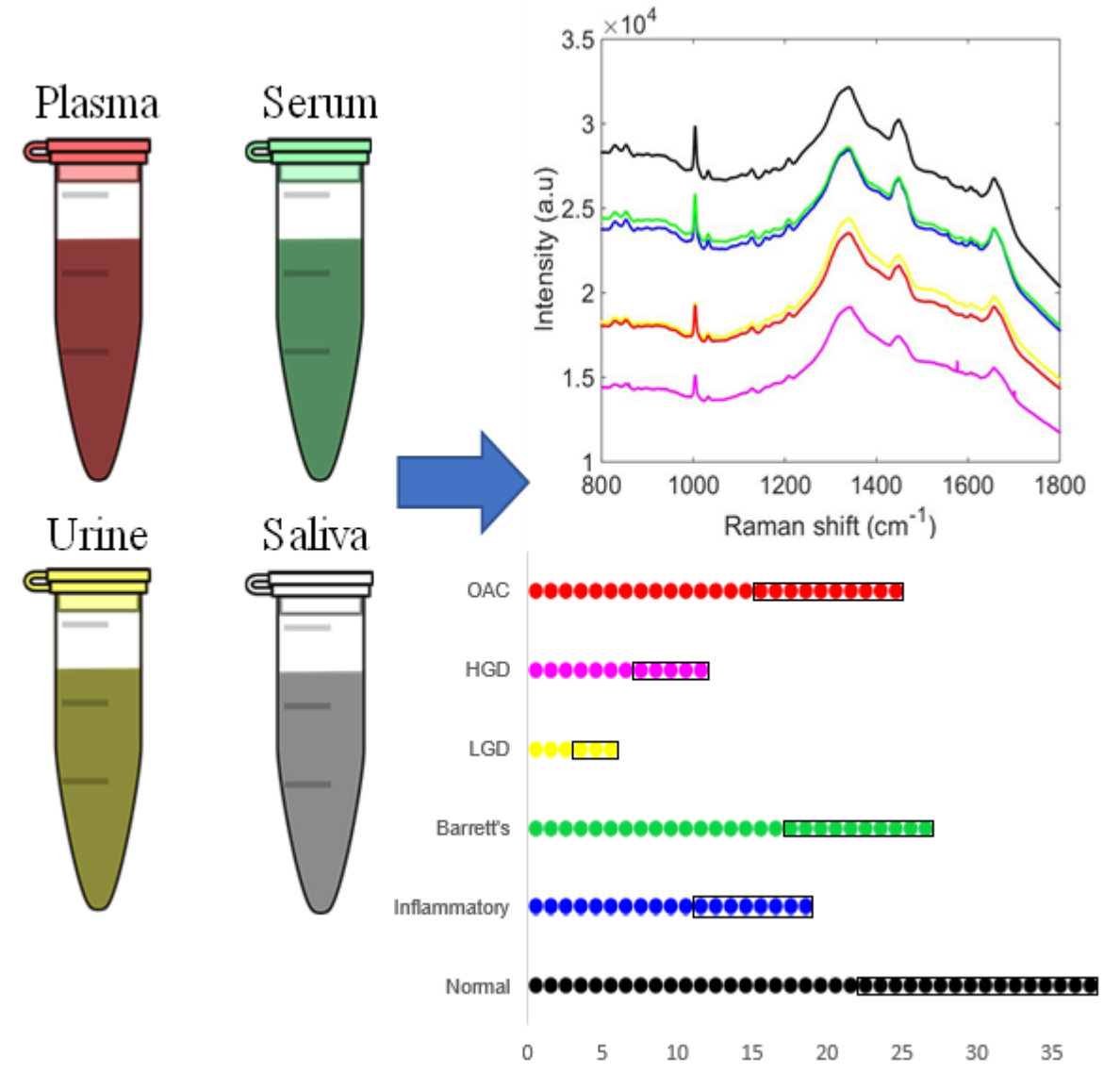
Raman spectroscopy

Raman point spectra acquisition was performed with an InVia Renishaw Raman spectrometer coupled with a charge-coupled device (CCD) detector and a Leica microscope. A 200-mW laser diode was used at a wavelength of 785 nm with a grating of 1200 lines/mm. Exposure time was set at 10 s, with 5% laser power, and 2 accumulations at a spectral range between 2000 cm^{-1} – 400 cm^{-1} . Ten-point spectra were taken per sample using a 20× objective to focus the laser beam on the sample.

Figure 19: Renishaw 100 System



Figure 20: Graphical abstract demonstrating how the four biofluids (plasma, serum, urine and saliva) were measured through Raman spectroscopy in order to detect oesophageal transformation stages to adenocarcinoma



Data Analysis and Chemometric Methods

The data import, pre-treatment and construction of chemometric classification models (principal component analysis quadratic discriminant analysis [PCA-QDA], successive projections algorithm quadratic discriminant analysis [SPA-QDA] and genetic algorithm quadratic discriminant analysis (GA-QDA)) were implemented in MATLAB R2014a software (MathWorks, USA) by using the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., USA) and laboratory-made routines. The raw spectra were pre-processed by cutting between 1800 cm^{-1} and 800 cm^{-1} (939 wavenumbers at 4 cm^{-1} spectral resolution). Cosmic rays were corrected using the Renishaw WiRE software system, and baseline distortions due to fluorescence interference were corrected with the asymmetric least squares (ALS) baseline correction algorithm (Eilers, 2004). For the PCA-QDA, SPA-QDA and GA-QDA models, samples were divided into training (60%), validation (20%) and prediction sets (20%) by applying the classic Kennard–Stone (KS) uniform sampling algorithm (Kennard and Stone, 1969) to the spectra as shown in Table 10. The optimum number of variables for SPA-QDA and GA-QDA were performed with an average risk G of QDA misclassification. Such a cost function was calculated in the validation set as:

$$G = \frac{1}{N_V} \sum_{n=1}^{N_V} g_n \quad (1)$$

where N_V is the number of validation spectra and g_n is defined as:

$$g_n = \frac{r^2(x_n, m_{I(n)})}{\min_{I(m) \neq I(n)} r^2(x_n, m_{I(m)})} \quad (2)$$

where $I(n)$ is the index of the true class for the n th validation object x_n . In this definition, the numerator is the squared Mahalanobis distance between object x_n (of class index $I(n)$) and the sample mean $m_{I(n)}$ of its true class. The denominator in Eq. (2) corresponds to the squared Mahalanobis distance between object x_n and the centre of the closest wrong class, $m_{I(m)}$. The minimum value of the cost function (maximum fitness) will be achieved when the selected variables from the original data are as close as possible to its true class and more distance as possible from its wrong class according to the validation samples. The GA routine was carried out during 100 generations with 200 chromosomes each. Crossover and mutation probabilities were set to 60% and 1%, respectively. Moreover, the algorithm was repeated three times, starting from different random initial populations. The best solution of GA (in terms of the fitness value) resulting from the three realisations was employed.

The calculation of classification quality parameters is a recommended standard practice for test evaluation (Cheung *et al.*, 2011). For this study, measures of test accuracy, such as sensitivity (proportion of positive samples correctly identified), specificity (proportion of negative samples correctly identified) and F-score (measurement of the model accuracy), were utilised. These quality metrics were calculated as follows:

$$\text{Sensitivity (\%)} = \frac{\text{TP}}{\text{TP}+\text{FN}} \times 100 \quad (3)$$

$$\text{Specificity (\%)} = \frac{\text{TN}}{\text{TN}+\text{FP}} \times 100 \quad (4)$$

$$\text{F-score} = \frac{2 \times \text{SENS} \times \text{SPEC}}{\text{SENS} + \text{SPEC}} \quad (5)$$

where TP stands for true positives, TN for true negatives, FP for false positives and FN for false negatives. SENS stands for sensitivity and SPEC for specificity.

All selected wavenumbers obtained from SPA-QDA and GA-QDA for all oesophageal stages of disease (*i.e.*, normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC) were confirmed by a Student's *t*-test (two-tailed, 95% confidence interval).

Results

The number of training, validation and prediction specimens (or spectra) in each category for biofluids is summarised in Table 10.

Table 10: Number of training, validation and prediction samples in each category of biofluids

Category – biofluids	Training	Validation	Test
Normal – plasma	21	7	7
Normal – serum	22	7	7
Normal – saliva	21	7	7
Normal – urine	23	7	8
Inflammatory – plasma	11	3	4
Inflammatory – serum	12	3	4
Inflammatory – saliva	11	3	4
Inflammatory – urine	12	3	4
Barrett’s oesophagus– plasma	16	5	6
Barrett’s oesophagus – serum	17	5	6
Barrett’s oesophagus – saliva	16	5	5
Barrett’s oesophagus – urine	17	5	5
LGD – plasma	3	1	2
LGD – serum	3	1	2
LGD – saliva	2	1	2
LGD – urine	3	1	2
HGD – plasma	7	2	3
HGD – serum	7	2	3
HGD – saliva	7	2	3
HGD – urine	7	3	3
OAC – plasma	13	4	5
OAC – serum	13	5	5
OAC – saliva	14	5	5
OAC – urine	15	5	6

Plasma dataset. Fig. 21A shows the average raw Raman spectra derived from blood plasma for all groups (normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC), respectively. Raman spectra of all oesophageal sample stages were compared after cosmic rays and baseline correction. The shape and trend of the six groups were very similar: there are four main Raman peaks at the position around 1004 cm^{-1} , 1335 cm^{-1} , 1450 cm^{-1} and 1660 cm^{-1} after fluorescence background removal (Figure 21B). Electron-rich groups (*e.g.*, C=O, C=N, and C=C) are the major source of features in Raman spectroscopy (Movasaghi *et al.*, 2011), and many Raman peaks are caused by the same molecular functional group belonging to different biomolecules in the material (Stewart *et al.*, 2012). However, there are still some visible spectral differences, notably the bands at 1004 cm^{-1} (collagen), 1335 cm^{-1} (CH_3CH_2 wagging mode of collagen), 1450 cm^{-1} (methylene deformation) and 1660 cm^{-1} (Amide I). Most of these bands decrease in amplitude from the normal group to the OAC group. The difference between groups can be observed more clearly from the averaged pre-processed spectrum of each group (Figure 21B). After pre-processing of the spectral data, chemometric techniques (PCA-QDA, SPA-QDA and GA-QDA) were applied to systematically classify all groups based on their Raman spectra.

The classification of the six oesophageal stages was developed by discriminant analysis using the Raman spectra between 800 cm^{-1} and 1800 cm^{-1} . The GA-QDA model with only 16 variables (namely 884 cm^{-1} , 1188 cm^{-1} , 1206 cm^{-1} , 1235 cm^{-1} , 1296 cm^{-1} , 1307 cm^{-1} , 1365 cm^{-1} , 1383 cm^{-1} , 1402 cm^{-1} , 1440 cm^{-1} , 1461 cm^{-1} , 1608 cm^{-1} , 1641 cm^{-1} , 1656 cm^{-1} , 1715 cm^{-1} , and 1793 cm^{-1} ; Fig. 21C and 21D) was found to give the highest classification accuracy ($>91.3\%$) in comparison with the other methods (PCA-QDA and SPA-QDA) in plasma (3 errors training; 3 errors validation). The figures of merit in the test set for inflammatory and Barrett's oesophagus samples using GA-QDA was 100%. For the other classes, GA-QDA achieved accuracies, sensitivities and specificities above $>80\%$. The PCA-QDA model using four PC scores (90% of the variance for all classes) achieved 100% accuracy, sensitivity, specificity and F-scores for LGD and OAC classes (Table 11). The SPA-QDA model also achieved a considerable high accuracy in the classification of HGD (100%) when applied using 30 selected wavenumbers, as shown in Table 11. Table ‡S5 lists the selected wavenumbers obtained with the GA-QDA model applied to the plasma samples along their respective tentative biomolecular assignments.

Figure 21: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using plasma samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 16 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.

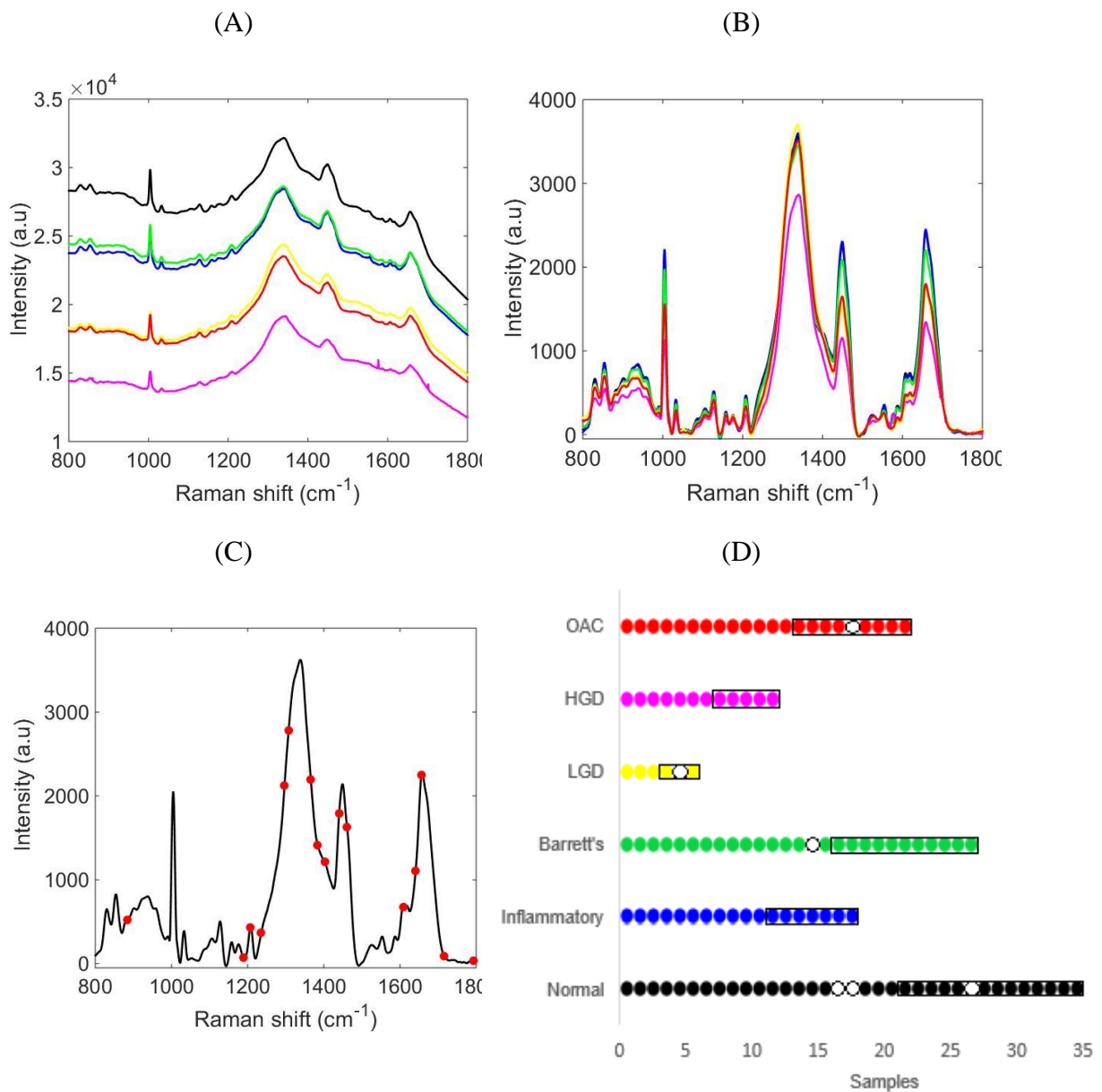


Table 11: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using plasma samples.

PCA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	91.3	91.3	95.6	100	95.6	100
Sensitivity (%)	100	90	100	100	95.2	100
Specificity (%)	71.4	100	80	100	100	100
F-Scores (%)	83.3	94.7	88.8	100	97.5	100
SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	82.6	86.9	95.6	95.6	100	86.4
Sensitivity (%)	87.5	90	94.4	100	100	94.4
Specificity (%)	71.4	66.6	100	100	100	60
F-Scores (%)	78.6	76.6	97.1	73.3	100	73.3
GA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	95.2	100	100	91.3	95.6	91.3
Sensitivity (%)	93.7	100	100	95.4	100	94.4
Specificity (%)	100	100	100	100	70	80
F-Scores (%)	96.7	100	100	83.3	66.6	86.6

Saliva dataset. Fig. 22A shows the average raw Raman spectra derived from saliva for all groups (normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC). Raman spectra of all stages of oesophageal disease were compared after cosmic rays and baseline correction. There are three main Raman peaks evident with the shape and trend similar for all groups: 1331 cm^{-1} , 1413 cm^{-1} and 1556 cm^{-1} in the raw spectra after fluorescence background removal (Figure 22B). In particular, strong peaks observed in the pre-processed spectra at 1336 cm^{-1} and 1664 cm^{-1} which indicate Amide III and CH₂ wagging vibrations from glycine backbone, and Amide I, respectively. These peaks are inherent to Raman spectra of saliva (Virkler and Lednev, 2010). Peaks at 852 cm^{-1} and 1128 cm^{-1} correspond to C-N stretching, CH₃ rocking and C-O vibrations, respectively.

The classification of the six oesophageal stages was developed by discriminant analysis using the Raman pre-processed spectra between 800 cm^{-1} and 1800 cm^{-1} . The classification for the test set using GA-QDA was equal to 100% for all FOM for all groups based on only 16 selected wavenumbers (namely 804 cm^{-1} , 848 cm^{-1} , 873 cm^{-1} , 943 cm^{-1} , 1012 cm^{-1} , 1020 cm^{-1} , 1091 cm^{-1} , 1163 cm^{-1} , 1198 cm^{-1} , 1326 cm^{-1} , 1397 cm^{-1} , 1404 cm^{-1} , 1453 cm^{-1} , 1528 cm^{-1} , 1552 cm^{-1} and 1765 cm^{-1}), as shown in Figure 22C and Figure 22D (4 errors training; 5 errors validation). The PCA-QDA model using seven PC scores (90% of the variance for all classes) achieved good results specifically for the inflammatory and HGD groups (>90% for all FOM) as can be seen in Table 12. The SPA-QDA model demonstrated excellent figures of merit in classifying OAC (100%) using 30 selected wavenumbers (Table 12). Table ‡S6 lists the selected wavenumbers obtained by GA-QDA for saliva samples with their tentative biomolecular assignments.

Figure 22: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using saliva samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 16 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. O - illustrates a misclassification.

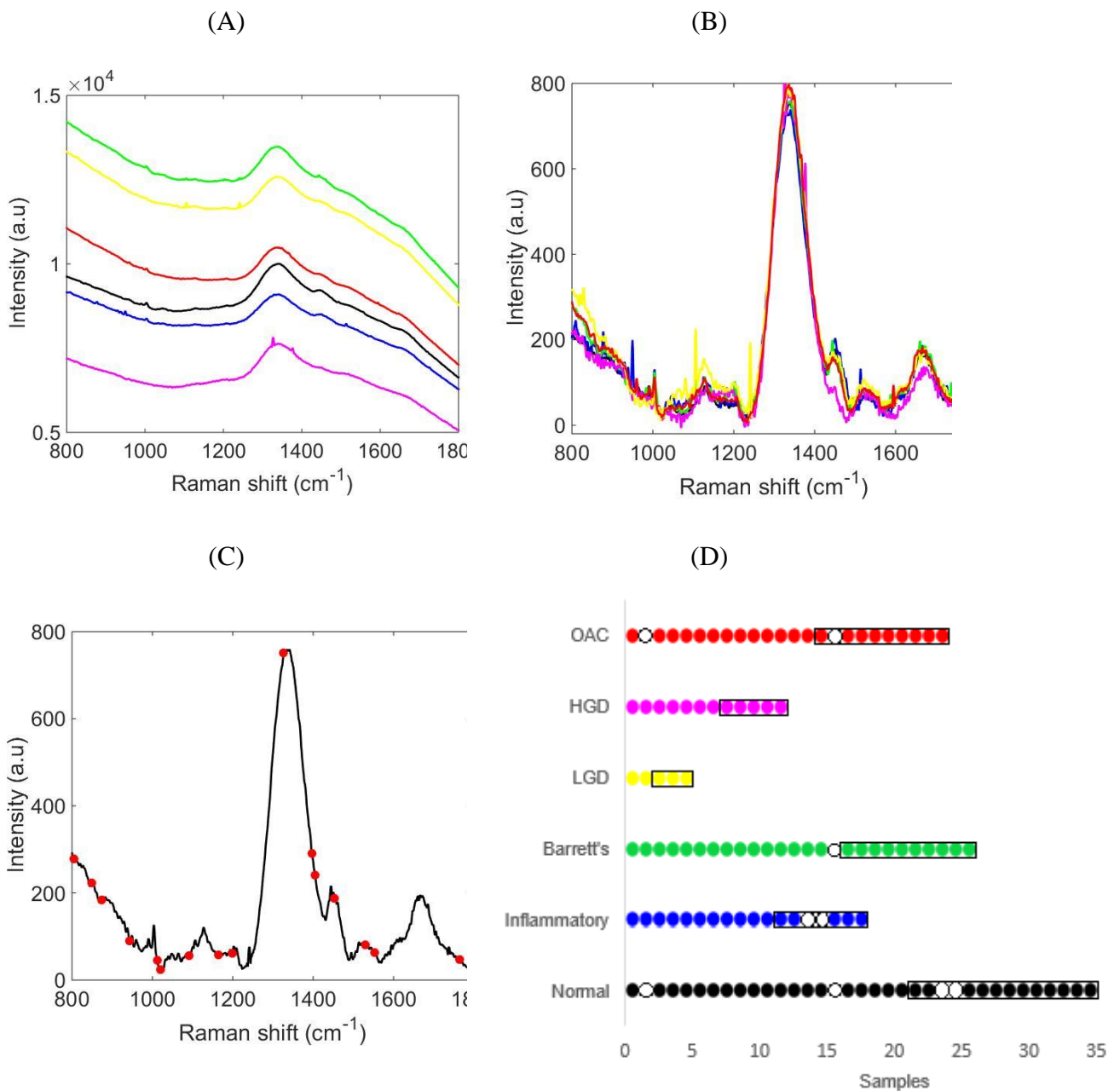


Table 12: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using saliva samples.

PCA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	82.6	91.3	95.6	95.6	95.6	86.9
Sensitivity (%)	87.5	90	100	100	95.2	94.4
Specificity (%)	71.4	100	80	0	100	60
F-Scores (%)	78.6	94.7	88.8	0	97.5	73.3
SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	95.6	95.6	95.6	95.6	91.3	100
Sensitivity (%)	100	95	100	95.4	95.2	100
Specificity (%)	85.7	100	80	100	50	100
F-Scores (%)	92.3	97.4	88.8	97.6	65.5	100

Serum Dataset. Fig. 23A shows the average raw Raman spectra derived from serum for all groups (normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC). There are three main Raman peaks with the shape and trend similar for all groups: 1327 cm^{-1} , 1443 cm^{-1} and 1662 cm^{-1} in the raw spectra after fluorescence background removal (Figure 23B). In particular, strong peaks observed in the pre-processed spectra at 1004 cm^{-1} , 1337 cm^{-1} , 1450 cm^{-1} and 1657 cm^{-1} indicate phenylalanine (proteins), Amide III, CH_2 bending and Amide I, respectively.

The PCA-QDA model using seven PC scores (95% of the variance for all classes) achieved good results specifically for inflammatory, HGD and OAC groups ($> 80\%$ FOM) as can be seen in Table 13. The SPA-QDA model demonstrated excellent figures of merit in classifying OAC (100%) using 30 selected wavenumbers. The classification on the test set using GA-QDA was 100% for all FOM for the normal and HGD groups based on 30 selected wavenumbers (namely, 821 cm^{-1} , 842 cm^{-1} , 894 cm^{-1} , 962 cm^{-1} , 989 cm^{-1} , 1035 cm^{-1} , 1042 cm^{-1} , 1047 cm^{-1} , 1062 cm^{-1} , 1132 cm^{-1} , 1146 cm^{-1} , 1162 cm^{-1} , 1165 cm^{-1} , 1249 cm^{-1} , 1279 cm^{-1} , 1282 cm^{-1} , 1321 cm^{-1} , 1362 cm^{-1} , 1402 cm^{-1} , 1414 cm^{-1} , 1415 cm^{-1} , 1450 cm^{-1} , 1471 cm^{-1} , 1550 cm^{-1} , 1688 cm^{-1} , 1711 cm^{-1} , 1726 cm^{-1} , 1727 cm^{-1} , 1731 cm^{-1} and 1789 cm^{-1}), as shown in Figure 23C. There were 2 errors in the training set and 1 error in the validation set. The selected wavenumbers by GA-QDA for serum with their respective tentative assignment are listed in Table ‡S7.

Figure 23: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using serum samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 30 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.

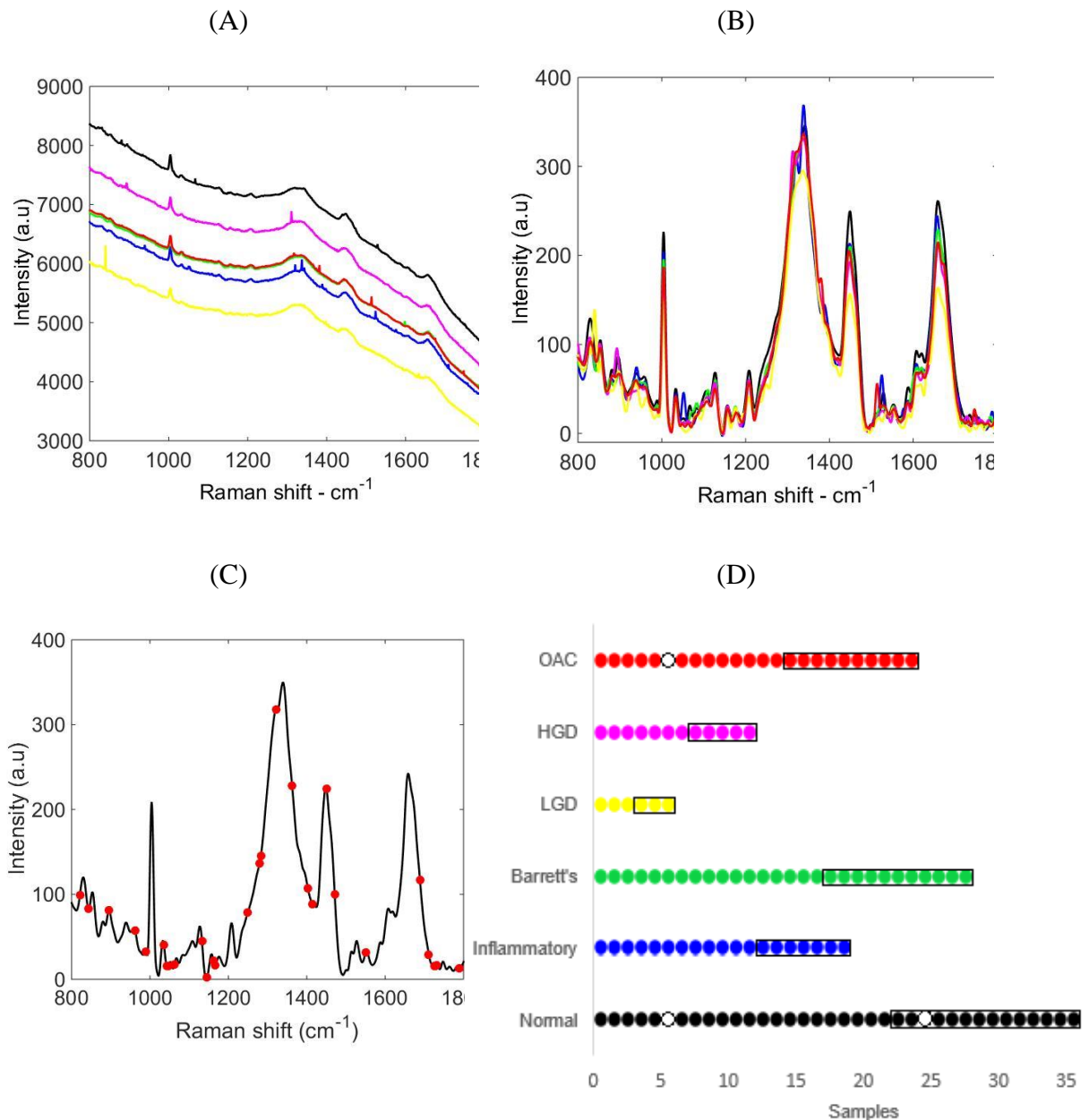


Table 13: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using serum samples.

PCA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	86.9	91.3	91.3	95.6	91.3	91.3
Sensitivity (%)	93.7	90.0	100	100	90.4	94.4
Specificity (%)	71.4	100	60	0	100	80.0
F-Scores (%)	81.0	94.7	75	0	95	86.6
SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	95.6	95.6	91.3	95.6	86.9	100
Sensitivity (%)	100	95.0	100	95.4	90.4	100
Specificity (%)	85.7	100	60	100	50	100
F-Scores (%)	92.3	97.4	75	97.6	64.4	100
GA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	100	95.6	95.6	95.6	100	95.6
Sensitivity (%)	100	100	94.4	100	100	94.4
Specificity (%)	100	66.6	100	60	100	100
F-Scores (%)	100	80.0	97.1	70	100	97.2

Urine Dataset. Fig. 24A shows the average raw Raman spectra derived from urine for all groups (normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC). The Raman spectra of all oesophageal disease stages were analysed after cosmic rays and baseline correction. There are two main Raman peaks with the shape and trend similar for all groups: 1012 cm^{-1} and 1340 cm^{-1} in the raw spectra after fluorescence background removal (Figure 24B). In particular, strong peaks are observed in the pre-processed spectra at 1012 cm^{-1} and 1336 cm^{-1} indicating C-O stretching in ribose and polynucleotide chain (DNA purine bases), respectively.

Classification of the six oesophageal stages was developed by discriminant analysis using the Raman spectra between 800 cm^{-1} and 1800 cm^{-1} . The GA-QDA model demonstrated excellent figures of merit (100%) for all stages to OAC based on only 29 selected wavenumbers (namely 845 cm^{-1} , 849 cm^{-1} , 858 cm^{-1} , 864 cm^{-1} , 877 cm^{-1} , 997 cm^{-1} , 1051 cm^{-1} , 1089 cm^{-1} , 1186 cm^{-1} , 1230 cm^{-1} , 1231 cm^{-1} , 1248 cm^{-1} , 1320 cm^{-1} , 1348 cm^{-1} , 1374 cm^{-1} , 1481 cm^{-1} , 1565 cm^{-1} , 1580 cm^{-1} , 1616 cm^{-1} , 1681 cm^{-1} , 1684 cm^{-1} , 1704 cm^{-1} , 1710 cm^{-1} , 1719 cm^{-1} , 1729 cm^{-1} , 1740 cm^{-1} , 1763 cm^{-1} and 1791 cm^{-1}), as can be seen in Figure 24C and 24D (no errors in training or validation sets). The PCA-QDA model using six PC scores (93% of the variance for all classes) achieved good results (>86% FOM) for the HGD group (Table 14). The SPA-QDA model achieved a considerable high FOM (100%) for four classes (normal, inflammatory, Barrett's oesophagus and HGD) using 30 selected wavenumbers (Table 14). Table 14 lists the selected wavenumbers obtained by GA-QDA for urine plasma samples with their respective tentative biomolecular assignments.

Figure 24: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using urine samples for Raman spectroscopy. The panel shows: (A) Average raw Raman spectrum in the region between 1800 cm^{-1} and 800 cm^{-1} ; (B) Average pre-processed Raman spectrum obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour); (C) 29 selected variables used by the GA-QDA model; (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.

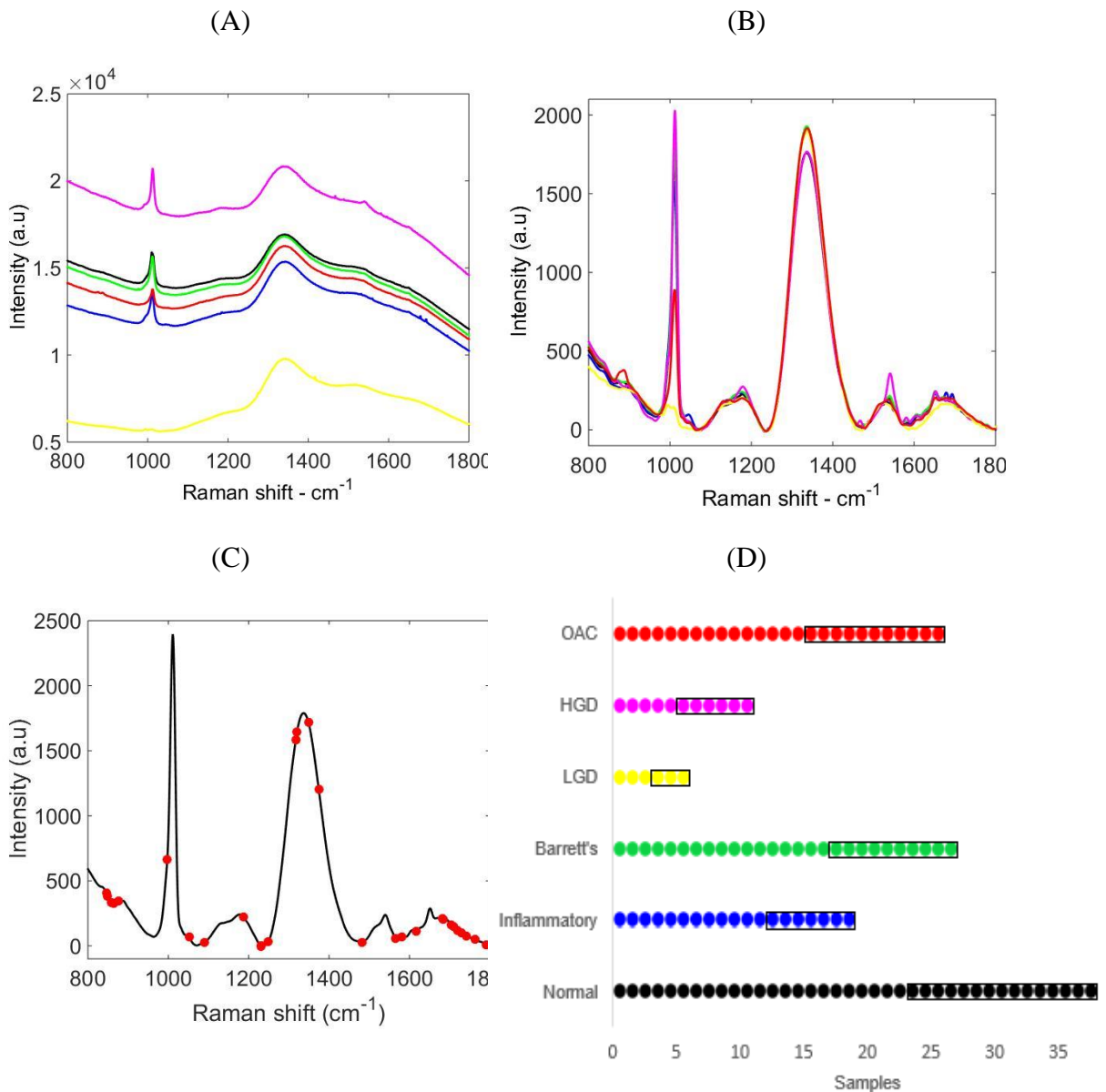


Table 14: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using urine samples.

PCA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	62.5	66.7	83.3	95.8	87.5	87.5
Sensitivity (%)	87.5	66.7	94.7	100	86.3	94.7
Specificity (%)	12.5	66.7	40	0	100	60
F-Scores (%)	21.8	66.7	56.2	0	92.7	73.4
SPA-QDA	Oesophageal stages					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	100	100	100	95.8	100	95.8
Sensitivity (%)	100	100	100	100	100	94.7
Specificity (%)	100	100	100	50	100	100
F-Scores (%)	100	100	100	50	100	97.3

Discussion

There have been no studies using Raman vibrational spectroscopy on biofluids in order to identify stages of oesophageal transformation to OAC. This study has demonstrated that Raman spectroscopy coupled multivariate classification techniques (PCA-QDA, SPA-QDA and GA-QDA) on biofluids can be used to identify oesophageal stages of disease to adenocarcinoma with excellent accuracy, sensitivity and specificity.

A key finding from this study suggests a reduction in some main biomolecules such as tryptophan (1365 cm^{-1}) and phenylalanine (1608 cm^{-1}) from the control group (patients with normal squamous epithelium) to the adenocarcinoma group in plasma. Aromatic amino acid metabolism may be dysfunctional in gastroesophageal cancer (Lai *et al.*, 2005). Patients with OAC have reduced plasma levels of 14 separate amino acids, including phenylalanine, and tryptophan (Lai *et al.*, 2005). This corresponds with our findings of reduced spectral scattering absorbed at these wavelengths in the OAC group.

In addition, spectral markers such as C-O stretching of ribose (1012 cm^{-1}), Amide III and CH₂ wagging vibrations from glycine backbone (1336 cm^{-1}), methylene deformation (1450 cm^{-1}), and Amide I (1660 cm^{-1}) were found as key discriminant features in saliva and urine between normal and progressive stages of oesophageal disease to OAC using the GA-QDA model. Ribose sugars are precursors to biosynthetic pathways generated by the Warburg effect, which are responsible for keeping cancer cells alive by generating energy through glycolysis, where glucose is converted to lactose for energy followed by lactate fermentation (Fadaka *et al.*, 2017). Amide III vibrations attributed to β -sheet and α -helix conformation in proteins are highly associated with oncogenesis (Abramczyk and Imiela, 2018; Talari *et al.*, 2019; Auner *et al.*, 2018). DNA methylation, which is an enzyme-induced chemical modification to the DNA structure where a methyl group is covalently bonded to the cytosine base is also involved in carcinogenesis (Wajed *et al.*, 2001). Amide I is known to be associated with cancer due to alterations in protein backbone conformation (Kuhar *et al.*, 2018).

Berger *et al.* (1999) initially introduced the idea that Raman had potential for the analysis of biofluids. Biofluid assays have numerous advantages including high accessibility. Sample processing is cheap and not laborious. This can be implemented in a clinical setting from routine investigations to intra-operative monitoring.

Raman spectroscopy has been performed on biofluid assays in a variety of cancer diagnostics. Harvey *et al.* (2008) conducted studies using Raman spectroscopy of urine to detect prostate cancer cells from bladder cancer cells. The authors identified that the Raman spectra suggested a higher concentration of nucleic acids and proteins in bladder cancer cells compared to the prostate cancer cells. Early work by Chan *et al.* (2006) focused on analysing white blood cells from plasma as a potential diagnostic tool for haematological malignancies such as Lymphoma and Leukaemia. The authors showed that single-cell Raman micro-spectroscopy was able to discriminate between normal human lymphocytes from transformed Jurkat (T cells) and Raji (B cells) lymphocyte cell lines based on highly reproducible biomolecular fingerprints. Multivariate statistical models based on the Raman spectra achieved a sensitivity of 98.3% for cancer detection, with 97.2% of the cells being correctly classified as belonging to the normal or transformed group (Chan *et al.*, 2006).

Eight biofluid samples (plasma $n = 2$; serum $n = 2$; saliva $n = 2$; urine $n = 2$) used for analysis from patients classified in OAC group had chemotherapy prior to biofluid collection. Previous studies have suggesting there are no significant Raman spectral differences associated with chemotherapy in biofluids (Feng *et al.*, 2010). The addition of these samples in the analysis would not statistically affect the average raw class spectral data for OAC for each biofluid and thus be insignificant with regards to the FOM created by the predictive chemometric models used.

Herein, the pre-processed Raman spectral datasets were analysed by classification methods based on QDA. QDA is a discriminant analysis algorithm based on a Mahalanobis distance (*i.e.*, distance between two points in a multivariate space) calculation that uses a separate variance-covariance matrix for each class (Morais and Lima, 2018). This increases the discrimination accuracy in complex biological mediums where classes having different variant structures are present (Morais and Lima, 2018). PCA-QDA, SPA-QDA and GA-QDA were used on the six groups of samples for all biofluids separately. GA-QDA was found to be very effective in the discrimination between all groups, since this algorithm achieved a high-quality performance rate using fewer wavenumbers (< 30 is an acceptable chemometric marker for class discrimination). The GA-QDA model obtained an accuracy of 100% for saliva and urine in all groups. Although the number of samples in the LGD and HGD groups are small, which is a limitation towards the predictive ability of this classifier, the overall results demonstrate promising evidence that Raman spectroscopy coupled with chemometric techniques can be used to distinguish different stages of oesophageal disease to adenocarcinoma.

By using liquid samples, analysis would be faster since no drying time would be necessary and the level of discrimination could improve since relevant compounds might evaporate during the drying process. In addition, by using specially prepared slides (silver coated) for surface enhanced Raman spectroscopy the sensitivity of this method could be improved further due to the large magnification of the Raman signal for these samples (Graham and Faulds, 2009). These were not used in the study as the slides were deemed not cost-effective for the pilot study. This study shows the potential of Raman spectroscopy and chemometrics for detecting oesophageal stages of disease through to OAC based on biofluids analysis.

Conclusion

Raman spectroscopy is a highly informative, non-destructive and robust technique that has been limitedly employed in the field of oesophageal disease. The results of this study show that Raman spectroscopy coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA) result in a powerful alternative approach for detection of oesophageal stages of disease to OAC in biofluids with an excellent accuracy, specificity and sensitivity in saliva and urine using the GA-QDA model. This pilot study is pioneer; but further work in this field including large multicentre studies should be performed in the future to validate these results.

Chapter

5



Declaration of work

To Whom it May Concern,

Mr Ishaan Maitra helped design the study and isolated all histology specimens with the aid of Katherine L. Ashton from the Lancashire Teaching Hospitals NHS Foundation Trust histology bank. Sample processing, dewaxing and storage was kindly undertaken by Katherine L. Ashton. Specimens were checked and compared with H&E specimens by Dr Danielle Bury to establish regions for spectroscopic analysis.

Mr Ishaan Maitra processed the specimens using the Bruker TENSOR 27 FTIR spectrometer at the University of Central Lancashire Research laboratory. Data was analysed by Camilo L.M. Morais and Kassio M.G. Lima. Ishaan Maitra, Camilo L.M. Morais and Kassio M.G. Lima prepared a paper for publication of the results with the support of Professor Francis L. Martin.

Signed

.....
Professor F L Martin

.....
Mr I Maitra

Attenuated total reflection Fourier-transform infrared spectral discrimination in human tissue of oesophageal transformation to adenocarcinoma

Introduction

Fourier-transform infrared (FTIR) spectroscopy has been employed to study cancer in gastrointestinal tissues including stomach (Li *et al.*, 2005) and intestinal tissue (Sahu *et al.*, 2010; Sahu *et al.*, 2017). Literature has found molecular abnormalities which occur before the change in morphology seen under the light microscope, as well significant differences between the spectra of malignant and corresponding normal tissues (Xu *et al.*, 2000; Wang *et al.*, 2003). The technology is capable of differentiating between unaffected and malignant tissues by comparing spectra for change in an array of diagnostic bands arising from phosphate, C–O and CH stretching vibrational modes. Chemometric methods such as principal component analysis (PCA) and hierarchical clustering analysis (HCA) are commonly used to separate spectra of normal and neoplastic regions. Despite these technological improvements, biomedical applications of FTIR-based analytical technique has not progressed. The majority of studies in this field have focused on *ex vivo* analysis.

This large *ex vivo* study including all classifications to OAC adds further validation to previous human studies identifying the potential for ATR-FTIR spectroscopy in differentiating between all classes in oesophageal transformation to OAC (normal; inflammatory; Barrett's oesophagus; Low-Grade Dysplasia, LGD; High-Grade Dysplasia, HGD; and OAC).

Materials and Methods

Sample collection

Ethical approval was granted by the East of England - Cambridge Central Research Ethics Committee from 2015 (Archival gastro-intestinal tissue, blood, saliva and urine collection; REC reference: 18/EE/0069; IRAS project ID: 242639). Ethics was also granted from the University of Central Lancashire (STEMH 909 application). Patients were identified from upper GI multi-disciplinary team meetings and pathology hospital databases. Potential patients (n = 120) were identified prospectively and consent for tissue was taken between October 2017 and June 2019 in a clinic or endoscopy setting. Samples had been coded appropriately as normal squamous epithelium, squamous epithelium with inflammation, intestinal metaplasia (Barrett's oesophagus), Low grade dysplasia (LGD), high grade dysplasia (HGD) and oesophageal adenocarcinoma (OAC). The archived samples had been routinely processed at the time of their acquisition and kept in paraffin embedded tissue blocks in ideal conditions as per local departmental protocols. Tissue blocks are embedded in paraffin wax as this ensures durability for long term storage without deterioration of the tissue sample. Paraffin embedding has been shown to affect bands at 1465 cm^{-1} (associated with aromatic structure – Wiens *et al.*, 2007) in the fingerprint region; and bands at 2850 cm^{-1} , 2918 cm^{-1} and 2956 cm^{-1} (Meuse and Barker, 2009). Deparaffinisation was hence performed to reduce spectral interference after cutting prior to commencing ATR-FTIR measurements using local hospital protocols using xylene and ethanol.

Contiguous sections with $4\text{ }\mu\text{m}$ thickness were prepared on FisherBrand™ slides. Contiguous samples are used so that each section closely resembles the other sections, thus, ensuring correlation between histology and spectroscopic measurements. A separate Consultant Histopathologist identified sections of the cut biopsies for an overall representative analysis of the tissue to be conducted. This was to ensure that spectral measurements would be taken from the appropriate area, and from the same area to avoid heterogeneity. British Society of Gastroenterology guidelines (Fitzgerald *et al.*, 2014) state that all cases of suspected dysplasia are to be reviewed by a second GI pathologist. All slides were left to dry prior to transportation in wooden slide boxes to the spectroscopy laboratory for analysis. All the samples were stored in a de-humidified glass container to prevent condensation and physical damage.

ATR-FTIR spectroscopy

Spectroscopic interrogation of tissue samples was performed at the Biomedical research laboratory at the University of Central Lancashire (UK). Histological diagnoses were unknown to those who performed IR spectroscopy. IR spectra were obtained using a Bruker TENSOR 27 FTIR spectrometer with a Helios ATR attachment containing a diamond crystal (Bruker Optics Ltd, Coventry, UK) and operated using OPUS 6.5 software. Spectra were acquired from 10 independent sample locations. Data acquisition parameters were: 8 cm^{-1} spectral resolution giving 4 cm^{-1} data spacing, 32 scans, 6 mm aperture setting and $2\times$ zero-filling factors. These standard settings are the optimum for subcellular interrogation. The ATR diamond crystal was washed with distilled water and dried with tissue paper between each sample and before each new slide. These are standard settings for optimum subcellular interrogation (Baker *et al.*, 2014). A background absorption spectrum was taken prior to each new sample for atmospheric correction. Average spectral point measurement tissue areas were between 2mm x 2mm and 4mm x 4mm.

Data analysis and chemometric methods

The data import, pre-treatment and construction of chemometric classification models were implemented in MATLAB R2014a software (MathWorks, USA) by using the PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., USA) and custom-made routines. Raw spectra were pre-processed by cutting between 1800 cm^{-1} and 900 cm^{-1} (235 wavenumbers), followed by rubberband baseline-correction and normalisation to the Amide I peak (1650 cm^{-1}). Before constructing the multivariate classification models (principal component analysis quadratic discriminant analysis – PCA-QDA, successive projections algorithm quadratic discriminant analysis – SPA-QDA, genetic algorithm quadratic discriminant analysis – GA-QDA) the samples were divided into training (60%), validation (20%) and prediction (20%) sets by the classic Kennard–Stone (KS) (Kennard and Stone, 1969) uniform sampling algorithm applied to the pre-processed IR spectra. The training samples were used in the modelling procedure, whereas the prediction set was only used in the final classification evaluation. The optimum number of variables for SPA-QDA and GA-QDA was determined according to an average risk G of misclassification. Such a cost function is calculated in the validation set as:

$$G = \frac{1}{N_V} \sum_{n=1}^{N_V} g_n \quad (1)$$

where N_V is the number of validation spectra and g_n is defined as:

$$g_n = \frac{r^2(x_n, m_{I(n)})}{\min_{I(m) \neq I(n)} r^2(x_n, m_{I(m)})} \quad (2)$$

In this definition, the numerator is the squared Mahalanobis distance between object x_n (of class index $I(n)$) and the sample mean $m_{I(n)}$ of its true class. The denominator in Eq. (2) corresponds to the squared Mahalanobis distance between object x_n and the centre of the closest wrong class, $m_{I(m)}$. The minimum value of the cost function (maximum fitness) will be achieved when the selected variables from the original data are as close as possible to its true class and more distance as possible from its wrong class according to the validation samples. The GA routine was carried out during 100 generations with 200 chromosomes each. Crossover and mutation probabilities were set to 60% and 1%, respectively. Moreover, the algorithm was repeated three times, starting from different random initial populations. The best solution (in terms of the fitness value) resulting from the three realisations of the GA was employed. The QDA classification score (Q_{ik}) is estimated using the variance-covariance matrix for each class k and an additional natural logarithm term, as follows:

$$Q_{ik} = (x_i - \bar{x}_k)^T \Sigma_k^{-1} (x_i - \bar{x}_k) + \log_e |\Sigma_k| - 2 \log_e \pi_k \quad (3)$$

where x_i is a vector with the input variables for sample i ; \bar{x}_k is the mean vector of class k ; Σ_k is the variance-covariance matrix of class k ; $\log_e |\Sigma_k|$ is the natural logarithm of the determinant of the variance-covariance matrix of class k ; and, $\log_e \pi_k$ is the natural logarithm of the prior probability term of class k . QDA forms a separated variance model for each class and does not assume that different classes have similar variance-covariance matrices, different to what is assumed by linear discriminant analysis (LDA) (Dixon and Brereton, 2009).

The calculation of figures of merit is a recommended standard practice to test model performance (Cheung *et al.* 2011; Eilers, 2004). Herein, measures of test accuracy including sensitivity (portion of positive samples correctly classified), specificity (portion of negative samples correctly classified), and F-score, which is a general measurement of the model accuracy, were performed. These quality metrics are calculated using the following equations:

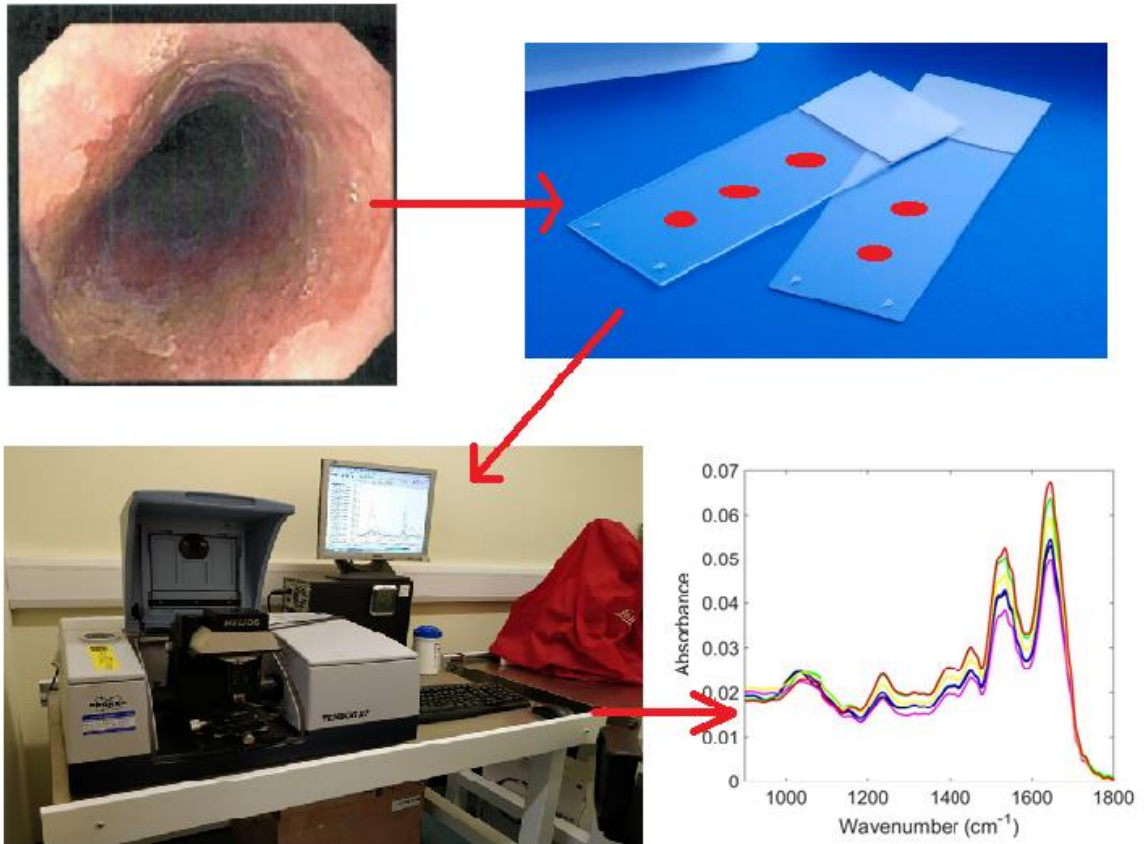
$$\text{Sensitivity (\%)} = \frac{\text{TP}}{\text{TP}+\text{FN}} \times 100 \quad (4)$$

$$\text{Specificity (\%)} = \frac{\text{TN}}{\text{TN}+\text{FP}} \times 100 \quad (5)$$

$$\text{F-score} = \frac{2 \times \text{SENS} \times \text{SPEC}}{\text{SENS} + \text{SPEC}} \quad (6)$$

where TP stands for true positives, TN for true negatives, FP for false positives and FN for false negatives. SENS stands for sensitivity and SPEC for specificity.

Figure 25: Graphical abstract demonstrating how oesophageal tissue is processed through ATR-FTIR spectroscopy in order to detect oesophageal transformation stages to adenocarcinoma



Results

The tissue specimens were categorised as follows: $n = 35$ normal, $n = 18$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD, and $n = 22$ OAC. Fig. 26A shows the average raw IR spectra derived from oesophageal tissue for the six oesophageal stages (normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC). Overall, the IR spectra for oesophageal stages appear to overlap in the biochemical fingerprint region (1800 cm^{-1} to 900 cm^{-1}), making it difficult to distinguish any subtle but significant differences visually. On closer analysis, notable distinguishing peaks that represent protein functional groups could be observed around 1650 cm^{-1} (Amide I) and 1550 cm^{-1} (Amide II). In addition, peaks were observed in the region of 1450 cm^{-1} - 1400 cm^{-1} (methylene deformation) and 1250 cm^{-1} - 1200 cm^{-1} (phosphate I). The major peaks of methyl groups of lipids and proteins could be found at around 1400 cm^{-1} (Amide III), 1225 cm^{-1} (asymmetric phosphate stretching vibrations, $\nu_{\text{as}}\text{PO}_2^-$) and 1080 cm^{-1} (symmetric phosphate stretching vibrations, $\nu_{\text{s}}\text{PO}_2^-$) [Table ‡S9]. To discriminate the six oesophageal stages, the spectral dataset was pre-processed using baseline correction and normalisation to the Amide I peak (Figure 26B). Average IR pre-processed spectra appear to overlap in the biochemical fingerprint region (1800 cm^{-1} to 900 cm^{-1}) making spectral observation and the discovery of spectral markers or signatures difficult. Therefore, chemometric techniques (PCA-QDA, SPA-QDA and GA-QDA) were adopted to classify normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC based on their IR spectra.

Figure 26C shows the wavenumbers associated with class differences. Fig. 26D shows the predicted class achieved by the GA-QDA model for all classes based on 16 selected wavenumbers (only 3 errors in the training set and 1 error in the prediction set). The GA-QDA model demonstrated good figures of merit for classifications to OAC except poor specificity and F-scores (50%) for LGD and HGD. The PCA-QDA model using the scores on seven PCs (90% of the total data variance) demonstrated an accuracy $>82\%$ for all classes (96.2% for inflammatory); a sensitivity $>82\%$ for all classes (100% for intestinal metaplasia); a specificity of $>40\%$ for all categories (100% for inflammatory) and F-scores $>50\%$ for all classes (97.4% for inflammatory) [Table 15]. SPA-QDA achieved excellent figures of merit (100%) for all classes to OAC except LGD and HGD using 7 wavenumbers (namely, 1392 cm^{-1} , 1485 cm^{-1} , 1539 cm^{-1} , 1585 cm^{-1} , 1624 cm^{-1} , 1643 cm^{-1} , and 1681 cm^{-1}).

Figure 26: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using tissue samples for ATR spectroscopy. The panel shows: (A) Average raw spectra in the ATR region of 1800 cm^{-1} to 900 cm^{-1} and (B) Average preprocessed ATR spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 16 selected variables used by GA-QDA model. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.

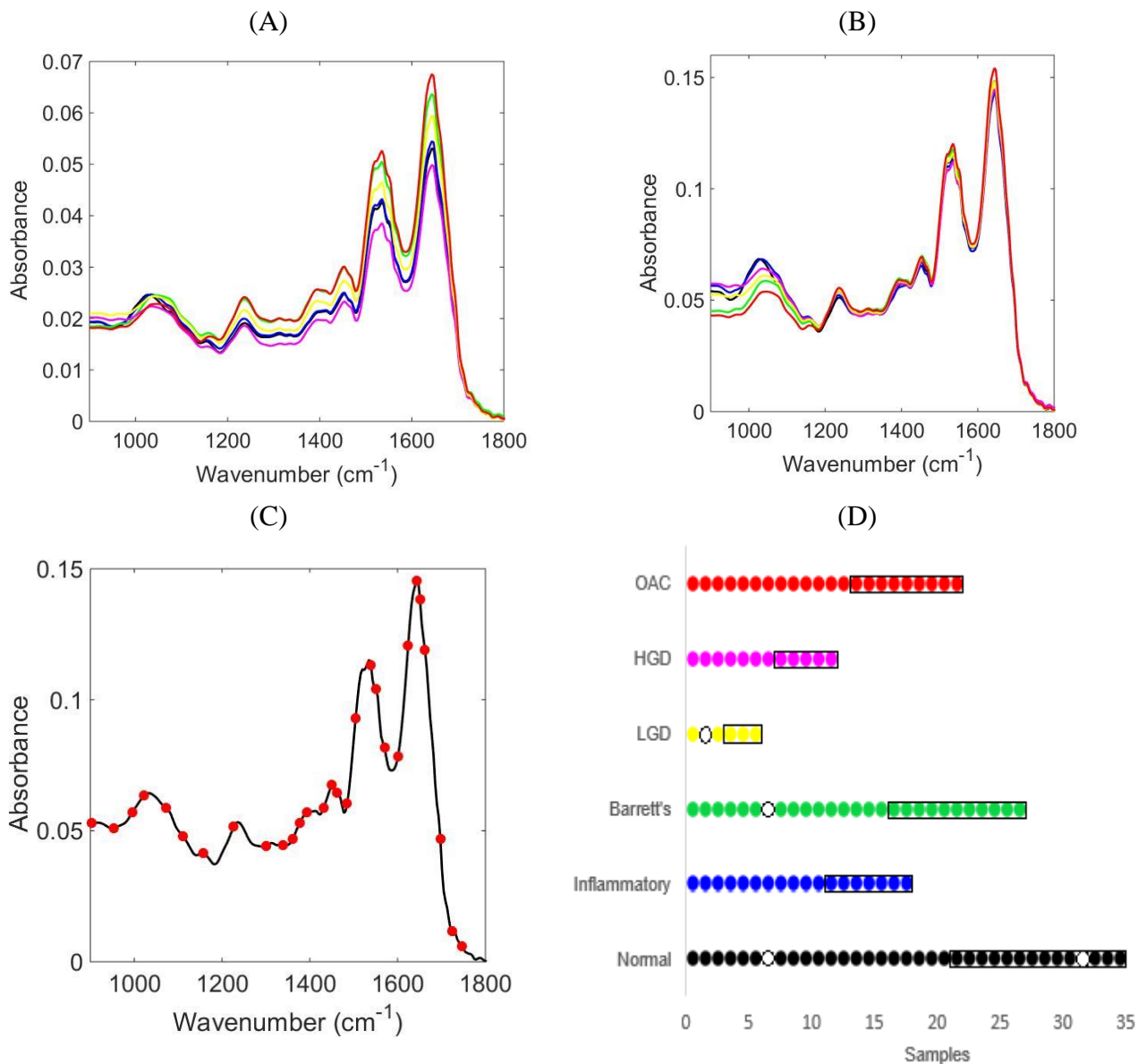


Table 15: Figures of merit (FOM) (accuracy, sensitivity, specificity and F-scores) and multivariate classification methods (PCA-QDA, SPA-QDA and GA-QDA) results for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using tissue samples.

FOM	PCA-QDA					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	82.6	95.6	91.3	91.3	86.9	82.6
Sensitivity (%)	82.3	95.0	100	95.4	90.4	94.4
Specificity (%)	83.3	100.0	66.6	50.0	50.0	40.0
F-Scores (%)	82.8	97.4	80	50.0	64.4	56.2
FOM	SPA-QDA					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	100.0	100.0	100.0	91.3	91.3	100.0
Sensitivity (%)	100.0	100.0	100.0	90.9	100	100.0
Specificity (%)	100.0	100.0	100.0	100	60.0	100.0
F-Scores (%)	100.0	100.0	100.0	95.2	60.0	100.0
FOM	GA-QDA					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	91.3	91.3	100.0	91.3	95.6	95.6
Sensitivity (%)	100.0	90.0	100.0	95.4	100.0	94.4
Specificity (%)	66.6	100.0	100.0	50	50.0	100.0
F-Scores (%)	80.0	94.7	100.0	50	66.6	97.1

Discussion

Initial research on ATR-FTIR spectroscopic analysis on human oesophageal tissue was conducted by Maziak *et al.* (2007). The authors performed an exploratory study in diagnosing OAC using FTIR. They found an increase in the nuclei-to-cytoplasm ratio, an enhancement in the phosphorylation of proteins, a decrease in the glycogen level, an increase in the distribution of protein segments, and an accumulation of triglycerides from OAC resected specimens compared to those with squamous epithelial specimens. Wang *et al.* (2003) identified significant differences between normal and malignant oesophageal tissues caused by the changes of content and space array of proteins, nucleic acids, sugars and fats in cells. Protein bands of Amide I and II were weak and broad in malignant tissues and strong and sharp in normal tissues.

Preliminary work on specialised intestinal metaplasia in Barrett's oesophagus as a precursor to OAC using FTIR was performed by Quaroni and Casson (2009). The authors found that specialised intestinal metaplasia had characteristic regions which displayed IR spectra with defined absorption features characteristic of glycoproteins. They added that highly fragmented regions identified in OAC likely reflected tumour heterogeneity. Wang *et al.* (2007) interrogated the use of FTIR to detect dysplastic mucosa in specialised intestinal metaplasia. The authors were able to demonstrate a sensitivity, specificity, positive predictive value, and total accuracy of 92%, 80%, 92% and 89%, respectively using FTIR to categorise dysplasia in 38 specimens. They concluded that this led to a better interobserver agreement between two gastrointestinal pathologists for dysplasia ($\kappa = 0.72$) vs. histology alone ($\kappa = 0.52$). Old *et al.* (2017) continued key research in this field mapping 22 oesophageal tissue samples from 19 patients. Key biochemical differences were identified by their spectral signatures: high glycogen content was seen in normal squamous tissue, high glycoprotein content was observed in glandular Barrett's oesophageal tissue, and high DNA content in dysplasia/adenocarcinoma samples. Classification of normal squamous samples vs. 'abnormal' samples (any stage of Barrett's oesophagus) was performed with 100% sensitivity and specificity. Neoplastic Barrett's oesophagus (dysplasia or adenocarcinoma) was identified with 95.6% sensitivity and 86.4% specificity. The authors concluded that highly accurate pathological classification can be achieved with FTIR measurement of frozen tissue sections in a clinically applicable time-frame. (Old *et al.*, 2017)

The findings from this study supports the use of ATR-FTIR coupled multivariate classification techniques (PCA-QDA, SPA-QDA and GA-QDA) in identifying oesophageal stages of disease to adenocarcinoma in human tissue. Excellent accuracy, sensitivity and specificity were achieved using different computational algorithms. This would encourage the investigation of screening for other cancers with known markers in an *ex vivo* setting. Classification of normal squamous samples *vs.* abnormal samples (any stage of Barrett's oesophagus to OAC) was performed with > 90% accuracy using the SPA-QDA model. This is comparable with previous study data (Old *et al.*, 2017; Wang *et al.*, 2007). Furthermore, this model demonstrated promise in defining LGD with all figures of merit >90.9%. This model however, revealed modest categorisation of HGD with regards to specificity (60%). LGD is difficult to distinguish from inflammation histologically due to subtle cellular differences. The overall risk of progression of LGD to HGD is approximately 9% (Duits *et al.*, 2019). This highlights the need for a clear method of categorisation between these groups given the large inter-observer availability in these groups. Further multicentre validation is necessary if *ex vivo* IR spectroscopy can be utilised as an adjunct to aid histopathologists in dysplasia diagnoses.

Conclusion

The results of this study show that ATR-FTIR spectroscopy of oesophageal tissue coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA) result in a powerful alternative approach for detection of oesophageal stages of disease to OAC in a fast, non-destructive fashion. The SPA-QDA model was particularly useful in its ability to define LGD (figures of merit > 90%), the only recognised marker for progression to OAC (Stone *et al.*, 2004).

Chapter

6

uclan



Declaration of work

To Whom it May Concern,

Mr Ishaan Maitra helped design the study and isolated all histology specimens with the aid of Katherine L. Ashton from the Lancashire Teaching Hospitals NHS Foundation Trust histology bank. Sample processing, dewaxing and storage was kindly undertaken by Katherine L. Ashton. Specimens were checked and compared with H&E specimens by Dr Danielle Bury to establish regions for spectroscopic analysis.

Mr Ishaan Maitra processed the specimens using the Renishaw 1000 system Raman spectrometer at the University of Central Lancashire Research laboratory. Data was analysed by Camilo L.M. Morais and Kassio M.G. Lima. Ishaan Maitra, Camilo L.M. Morais and Kassio M.G. Lima prepared a paper for publication of the results with the support of Professor Francis L. Martin.

Signed

.....
Professor F L Martin

.....
Mr I Maitra

Discrimination of oesophageal transformation to adenocarcinoma on human tissue samples using Raman microscopy

Introduction

An approach to oesophageal cancer screening in the general population based on tissue interrogated by Raman spectroscopy linked with variable selected methods for classification could be the potential to segregate stages to oesophageal adenocarcinoma.

Raman spectroscopy to delineate benign oesophageal diseases is in its infancy. Small volume studies have been performed in this field. One such study by Hiremath *et al.* (2017) interrogated 11 oesophageal specimens: healthy ($n = 5$), GORD ($n = 3$) and eosinophilic oesophagitis ($n = 3$). The authors established discriminant Raman peaks associated with Amide I, DNA, lipids, phospholipids and phenylalanine in eosinophilic oesophagitis in comparison to GORD and normal specimens. Other pertinent findings were a sensitivity of 81%, a specificity of 75% and an overall accuracy 75% when comparing healthy specimens to diseased ones (eosinophilic oesophagitis and GORD).

An *ex vivo* study by Chen *et al.* (2013) interrogated 54 oesophageal cancer tissues and 55 normal tissues in the $400\text{ cm}^{-1} - 1750\text{ cm}^{-1}$ range. The mean Raman spectra showed significant differences between the two groups. Tentative assignments of the Raman bands suggested some changes in protein structure, a decrease in the relative amount of lactose, and increases in the percentages of tryptophan, collagen and phenylalanine content in oesophageal cancer tissue compared to those in squamous tissue. The authors utilised diagnostic algorithms based on principal component analysis (PCA) and linear discriminate analysis (LDA). They achieved a diagnostic sensitivity of 87.0% and specificity of 70.9% for separating cancer from normal oesophageal tissue samples.

Several studies have investigated the ability of Raman to identify early stages of malignancy and thereby conceivably improve clinical outcomes. Kendall *et al.* (2003) interrogated 77 oesophageal specimens ($n = 27$ normal squamous epithelium; Barrett's mucosa [$n = 4$ cardiac-type Mucosa, $n = 8$ fundic-type mucosa, $n = 13$ intestinal metaplasia]; $n = 5$ low-grade dysplasia LGD, $n = 11$ high-grade dysplasia HGD, and $n = 9$ OAC). These classifications were simplified into 3 groups (normal, Barrett's oesophagus and OAC). The authors established that pre-malignant and malignant oesophageal tissues were associated with discriminant Raman peaks indicative of increased nucleic acids, reduced carotenoids, decreased glycogen and altered protein conformations when compared to the healthy

samples. Sensitivities and specificities were found to be 97% and 99% for normal, 84% and 98% for Barrett's oesophagus, and 94% and 93% for OAC.

Stone *et al.* (2004) analysed oesophageal biopsy samples from 44 patients. The authors interrogated 8 different pathology states (normal squamous epithelium, cardiac Barrett's tissue, fundic, Barrett's intestinal metaplasia, HGD; OAC; squamous dysplasia and squamous cell carcinoma). The authors concluded they were able to discriminate the columnar cell pathological groups from OAC with 84-97% sensitivity and 93-99% specificity. Additionally, when combining both columnar and squamous epithelial tissues, the authors were able to discriminate between both pathologies with 73-100% sensitivity and 92-100% specificity.

Hutchings *et al.* (2010) performed Raman microscopy on 49 patients categorising biopsy specimens into normal, low risk (Barrett's esophagus) and high risk (dysplasia and adenocarcinoma) groups. The authors found an overall validated classification model performance/accuracy of 97.7% (sensitivity 95 – 100%; specificity 98.6 – 100%) between the three groups.

Bergholt *et al.* (2011) conducted an *in vivo* study where a total of 75 oesophageal tissue sites from 27 patients were measured. An optical probe comprised a central fiber of 200 μm , for delivery of the laser light (785 nm wavelength) to the tissue, surrounded by thirty-two 200 μm collection fibers was introduced to establish real-time spectra. Forty-two *in vivo* Raman spectra were from normal tissues and 33 *in vivo* Raman spectra were from malignant tumors (adenocarcinoma $n = 27$, squamous cell carcinoma $n = 6$) as confirmed by histopathology OAC tissue showed distinct Raman signals associated with cell proliferation, lipid reduction, abnormal nuclear activity and neovascularisation. Using an LDA algorithm, the authors demonstrated an accuracy of 96% (*i.e.*, sensitivity of 97.0% and specificity of 95.2%) for a diagnosis of oesophageal malignancy. The group concluded that an image-guided Raman endoscopy technique in conjunction with biomolecular modeling has promising potential for the real-time, *in vivo* diagnosis and detection of oesophageal cancer during endoscopic examination.

The largest study to date interrogating Barrett's oesophagus specimens was performed Bergholt *et al.* (2014) including 373 patients subjected to multimodal real-time optical imaging. The authors focused on three groups (columnar lined oesophagus without goblet cells $n = 907$ spectra; nondysplastic Barrett's oesophagus $n = 318$ spectra; Barrett's positive for HGD $n = 177$ spectra). They found increased Raman signals associated with nucleic acids which indicated abnormal DNA content and a hyperchromatic state of neoplastic cells. Their methods generated 79% sensitivity and 74% specificity for the detection of dysplasia.

This study presents a large *ex vivo* study which adds further validation to previous human studies identifying the potential for Raman spectroscopy in differentiating between all classes in oesophageal transformation to OAC (normal; inflammatory; Barrett's oesophagus; Low-Grade Dysplasia, LGD; High-Grade Dysplasia, HGD; and OAC).

Material and Methods

Sample collection

Ethical approval was granted by the East of England - Cambridge Central Research Ethics Committee from 2015 (Archival gastro-intestinal tissue, blood, saliva and urine collection; REC reference: 18/EE/0069; IRAS project ID: 242639). Ethics was also granted from the University of Central Lancashire (STEMH 909 application). Patients were identified from upper gastro-intestinal (GI) multi-disciplinary team meetings and pathology hospital databases. Potential patients (n = 120) were identified prospectively and consent for tissue was taken between October 2017 and June 2019 in a clinic or endoscopy setting. Samples had been coded appropriately as normal squamous epithelium, squamous epithelium with inflammation, intestinal metaplasia (+/- dysplasia) or oesophageal adenocarcinoma. The archived samples had been routinely processed at the time of their acquisition and kept in paraffin embedded tissue blocks in ideal conditions as per local departmental protocols. Tissue blocks are embedded in paraffin wax as this ensures durability for long term storage without sample degradation. Paraffin embedding has been shown to affect 1465 cm^{-1} , 2850 cm^{-1} , 2918 cm^{-1} and 2956 cm^{-1} bands (Wiens *et al.*, 2007). Furthermore, paraffin wax produces a significant Raman signal with distinctive peaks (at 888 cm^{-1} , 1063 cm^{-1} , 1133 cm^{-1} , 1171 cm^{-1} , 1296 cm^{-1} , 1419 cm^{-1} , 1441 cm^{-1} and 1462 cm^{-1}) associated with aromatic structure. These peaks are in the fingerprint region and would render the interpretation of tissue spectra impossible (Mian *et al.*, 2014). De-paraffinisation was hence performed after cutting prior to commencing Raman measurements using local hospital protocols using xylene and ethanol.

Contiguous sections of $4\text{ }\mu\text{m}$ thickness were prepared on FisherBrand™ slides (Aluminum foil covered). Contiguous samples are used so that each section closely resembles the other sections, thus, ensuring correlation between histology and spectroscopic measurements. A separate Consultant Histopathologist identified sections of the cut biopsies for an overall accurate representative analytical study of the tissue. This was to ensure that spectral measurements would be taken from the appropriate area, and from the same area for the differing technologies to avoid heterogeneity in the cut tissue samples. British Society of Gastroenterology guidelines (Fitzgerald *et al.*, 2014) state that all cases of suspected dysplasia are to be reviewed by a second GI pathologist.

All slides were left to dry prior to transportation in wooden slide boxes to the Biomedical Research Laboratory for analysis. All of the samples were stored in a de-humidified glass container to prevent condensation and physical damage.

Raman spectroscopy

Raman point spectra were collected with an InVia Renishaw Raman spectrometer coupled with a charge-coupled device (CCD) detector and a Leica microscope at the Biomedical Research Laboratory at the University of Central Lancashire, UK. A 200-mW laser diode was used at a wavelength of 785 nm with a grating of 1200 lines/mm, and the system was calibrated to 520.5 cm^{-1} with a silicon source before every run. After trial-and-error measurements to optimise the experimental parameters, we utilised a 10 s exposure time, 5% laser power, and 2 accumulations at a spectral range $2000\text{ cm}^{-1} - 400\text{ cm}^{-1}$ to achieve optimum spectral quality. Ten random point-spectra were taken per sample using a 50× objective to focus the laser beam on the sample (Numerical aperture 0.75). For each sample, the H&E section was scanned, in high resolution, onto computer software (Microsoft PowerPoint 2018). Average spectral point measurement tissue areas were between 2mm x 2mm and 4mm x 4mm. This allowed the regions that best reflect the overall diagnosis to be highlighted and labelled. This is a standard of practice when using Raman spectroscopy to characterise biological materials (Butler *et al.*, 2016).

Data analysis and chemometric methods

The data import, pre-treatment and construction of chemometric classification models (PCA-QDA, SPA-QDA and GA-QDA) were implemented in MATLAB R2014a software (MathWorks, USA) by using PLS Toolbox version 7.9.3 (Eigenvector Research, Inc., USA) and laboratory-made routines. Raw spectra were pre-processed by cutting between 1800 cm^{-1} and 800 cm^{-1} (939 wavenumbers at 1 cm^{-1} spectral resolution). Cosmic rays were corrected using the Renishaw WiRe software. Baseline contribution due to fluorescence effect was corrected with the asymmetric Least Squares (ALS) algorithm (Eilers, 2004). For PCA-QDA, SPA-QDA and GA-QDA models, the samples were divided into training (60%), validation (20%) and prediction sets (20%) by applying the classic Kennard-Stone (Kennard and Stone, 1969) uniform sampling algorithm to the IR spectra. The leave-one-out cross-validation was implemented to avoid overfitting. The optimum number of variables for SPA-QDA and GA-QDA was performed with an average risk G of QDA misclassification. Such a cost function is calculated in the validation set as:

$$G = \frac{1}{N_V} \sum_{n=1}^{N_V} g_n \quad (1)$$

Where g_n is defined as:

$$g_n = \frac{r^2(x_n, m_{I(n)})}{\min_{I(m) \neq I(n)} r^2(x_n, m_{I(m)})} \quad (2)$$

Where I_n is the index of the true class for the n th validation object x_n .

In this definition, the numerator is the squared Mahalanobis distance between object x_n (of class index I_n) and the sample mean $m_{I(n)}$ of its true class. The denominator in Eq. (2) corresponds to the squared Mahalanobis distance between object x_n and the centre of the closest wrong class. The minimum value of the cost function (maximum fitness) will be achieved when the selected variables from the original data are closer as possible to its true class and more distance as possible from its wrong class according to the validation samples. The GA routine was carried out during 100 generations with 200 chromosomes each. Crossover and mutation probabilities were set to 60% and 1%, respectively. Moreover, the algorithm was repeated three times, starting from different random initial populations. The best solution (in terms of the fitness value) resulting from the three realisations of the GA was employed.

Figures of merit can be calculated in the prediction set to evaluate the model's predictive performance (Cheung *et al.*, 2011) For this study, measures of test accuracy, such as sensitivity (probability that a test result will be positive when the disease is present), specificity (probability that a test result will be negative when the disease is not present) and F-score (balanced measurement of the model accuracy) were performed. These quality metrics have been calculated using the following the equations:

$$\text{Sensitivity (\%)} = \frac{\text{TP}}{\text{TP}+\text{FN}} \times 100 \quad (3)$$

$$\text{Specificity (\%)} = \frac{\text{TN}}{\text{TN}+\text{FP}} \times 100 \quad (4)$$

$$\text{F-score} = \frac{2 \times \text{SENS} \times \text{SPEC}}{\text{SENS} + \text{SPEC}} \quad (5)$$

where SENS stands for sensitivity; SPEC stands for specificity; TP stands for true positives; TN stands for true negatives; FP stands for false positives; and FN stands for false negatives.

Results

The tissues were categorised as follows: The tissue specimens were categorised as follows: $n = 35$ normal, $n = 18$ inflammatory, $n = 27$ Barrett's oesophagus, $n = 6$ LGD, $n = 12$ HGD and $n = 22$ OAC.

Figure 27A shows the average raw Raman spectra derived from tissue for all groups (normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC). Raman spectra of all stages to OAC also were compared after baseline correction. There are six main Raman peaks with shape and trend similar for all groups: 1004 cm^{-1} , 1064 cm^{-1} , 1131 cm^{-1} , 1297 cm^{-1} , 1441 cm^{-1} and 1662 cm^{-1} after fluorescence background removal (Fig. 27B). In particular, discriminant strong peaks in the pre-processed spectra at 1296 cm^{-1} , 1448 cm^{-1} , and 1668 cm^{-1} which indicate CH₂ deformation, CH₂/CH₃ deformation and Amide I in proteins, respectively, are clearly observed. PCA-QDA, SPA-QDA and GA-QDA algorithms were applied to the pre-processed spectral data to systematically classify normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC.

The classification of the six oesophageal stages was developed by discriminant analysis using the Raman spectra between 1800 cm^{-1} and 800 cm^{-1} . The PCA-QDA model using seven PC scores (95% of the original data variance) achieved excellent classification results (100%) for LGD and OAC groups as demonstrated in Table 16. Excellent FOM were achieved using 30 selected wavenumbers in SPA-QDA for classification of HGD samples (100%). The SPA-QDA model demonstrated accuracy, sensitivity and specificity >90% for all classifications of oesophageal disease to OAC except HGD. The correct classification for validation models (test set) using GA-QDA achieved 100% for the inflammatory and Barrett's oesophagus groups based on 16 selected wavenumbers (800 cm^{-1} , 837 cm^{-1} , 897 cm^{-1} , 922 cm^{-1} , 997 cm^{-1} , 1183 cm^{-1} , 1267 cm^{-1} , 1290 cm^{-1} , 1301 cm^{-1} , 1578 cm^{-1} , 1585 cm^{-1} , 1608 cm^{-1} , 1704 cm^{-1} , 1712 cm^{-1} , 1776 cm^{-1} , 1800 cm^{-1}), as shown in Figure 27C and 27D (no errors in training or prediction sets). The GA-QDA model generated poor specificity for LGD (60.0%) and HGD (50.0%). The PCA-QDA model appears to be the optimum model giving more consistent results for all classes of disease to OAC (F-Scores ranging from 83.3–100%); thus, being a better classifier for clinical diagnosis. Table ‡S10 lists the selected wavenumbers obtained for GA-QDA models for tissue samples with their tentative biomolecular assignments.

Figure 27: Comparison of normal/inflammatory/Barrett's oesophagus/LGD/HGD/OAC oesophageal stages using tissue samples for Raman spectroscopy. The panel shows: (A) Average raw spectra in the Raman region between 1800 cm^{-1} - 800 cm^{-1} ; (B) Average preprocessed Raman spectra obtained from all stages segregated into normal (black colour) vs. inflammatory (blue colour) vs. Barrett's oesophagus (green colour) vs. LGD (yellow colour) vs. HGD (magenta colour) vs. OAC (red colour). (C) 16 selected variables by the GA-QDA model. (D) Predicted class vs. samples used for training and prediction sets (rectangular box), where each circle marker represents one sample for GA-QDA. \circ - illustrates a misclassification.

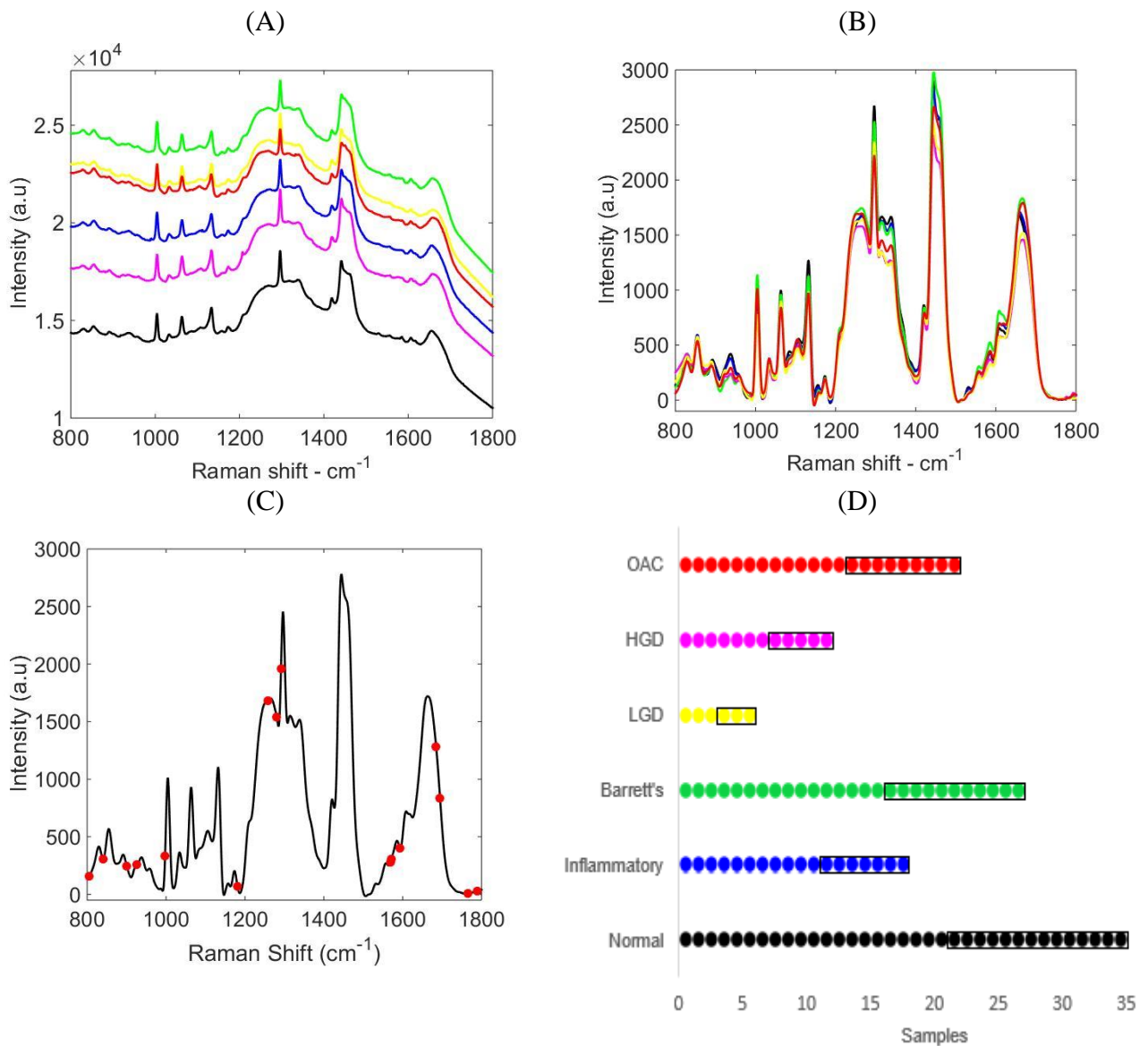


Table 16: Figures of merit (accuracy, sensitivity, specificity and F-scores) for the multivariate classification models (PCA-QDA, SPA-QDA and GA-QDA) to distinguish normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using tissue samples.

FOM	PCA-QDA					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	91.3	91.3	95.6	100	95.6	100
Sensitivity (%)	100	90.0	100	100	95.2	100
Specificity (%)	71.4	100	80.0	100	100	100
F-Scores (%)	83.3	94.7	88.8	100	88.8	100
FOM	SPA-QDA					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	82.6	86.9	95.6	95.6	100	86.9
Sensitivity (%)	87.5	90.0	94.4	100	100	94.4
Specificity (%)	71.4	66.6	100	50.0	100	60.0
F-Scores (%)	78.6	76.6	97.1	50.0	100	73.3
FOM	GA-QDA					
	<i>Normal</i>	<i>Inflammatory</i>	<i>Barrett's oesophagus</i>	<i>LGD</i>	<i>HGD</i>	<i>OAC</i>
Accuracy (%)	95.2	100	100	91.3	95.6	91.3
Sensitivity (%)	93.7	100	100	95.4	100	94.4
Specificity (%)	100	100	100	60.0	50.0	80
F-Scores (%)	96.7	100	100	60.0	66.6	86.6

Discussion

An approach to oesophageal cancer screening in the general population based on *in vivo* tissue interrogated by Raman spectroscopy linked with variable selected methods for classification could be the potential to segregate stages of oesophageal adenocarcinoma. Our group has discovered the use of Raman coupled multivariate classification techniques (PCA-QDA, SPA-QDA and GA-QDA) in identifying oesophageal stages of disease to adenocarcinoma has achieved excellent accuracy, sensitivity and specificity, encouraging investigation of screening for others cancers with known markers in an *ex vivo* setting. Discriminant strong peaks in the pre-processed spectra at 1296 cm^{-1} , 1448 cm^{-1} , and 1668 cm^{-1} which indicate CH₂ deformation, CH₂/CH₃ deformation and Amide I in proteins, respectively, are clearly observed. The PCA-QDA model appears to be the optimum model giving more consistent results for all classes of disease to OAC (F-Scores ranging from 83.3–100%); thus, being a better classifier for clinical diagnosis. The GA-QDA model was excellent in categorising Barrett's oesophagus.

From the existing literature, Raman spectroscopy has been applied for the diagnosis and identification of oesophageal pre-malignant and malignant conditions. *Ex vivo* studies all range from diagnostic accuracies for establishing non-dysplastic to dysplastic tissue from 88 – 97% (Kendall *et al.*, 2003; Stone *et al.*, 2004; Chen *et al.*, 2013; Hutchings *et al.*, 2010; Almond *et al.*, 2014). This compares with our group who has established an accuracy of 91% to 100% in classifying non-dysplastic from dysplastic tissue using the PCA-QDA model.

Peaks at 1296 cm^{-1} , 1448 cm^{-1} and 1668 cm^{-1} are clearly visible on the Raman profiles. Although the tissues samples underwent de-parrafinisation, these peaks may be caused due to residual contributions from paraffin wax. De-parrafinisation protocols with xylene washes have been found to be ineffective at the complete removal of wax. Subsequent de-waxing washes may still leave wax contributions at 1062 cm^{-1} , 1296 cm^{-1} , and 1441 cm^{-1} (Faoláin *et al.*, 2005). The units associated with spectroscopic intensity are arbitrary units (A.U.). The relative intensity to the baseline of each spectrum is relevant in this case. Hence, the spectra have been offset to appreciate clarity in the analysis.

The use of Raman spectroscopy in real-time endoscopy should not be undervalued. It is a commanding tool resulting in large volumes of data on the biochemical composition of tissues. However, with the current financial strain on the NHS, its application in the clinical setting needs to be carefully considered.

Conclusion

The results of this study show that Raman spectroscopy of oesophageal tissue coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA), result in a powerful alternative approach for detection of oesophageal stages of disease to OAC. This novel approach has potential to be used for early detection of neoplastic changes in susceptible epithelium with theoretical benefits for patient treatment in surveillance programs through the detection of premalignant conditions. Multicentre validation is necessary if *ex vivo* Raman spectroscopy can be utilised as an adjunct to aid histopathologists in dysplasia diagnoses.

Future studies are needed to focus on Raman spectroscopy in real-time endoscopy increasing its value as a tool for a 'one stop shop' service for patients. This may subsequently reduce the number of endoscopies patients undergo in the future thus being cost-effective in the difficult financial conditions currently experienced in the National Health Service.

Chapter

7

uclan



Declaration of work

To Whom it May Concern,

Mr Ishaan Maitra helped design the study and isolated all histology specimens with the aid of Katherine L. Ashton from the Lancashire Teaching Hospitals NHS Foundation Trust histology bank. Sample processing, dewaxing and storage was kindly undertaken by Katherine L. Ashton. Specimens were checked and compared with H&E specimens by Dr Danielle Bury to establish regions for spectroscopic analysis.

Mr Ishaan Maitra processed the specimens using the Renishaw 1000 system Raman spectrometer at the University of Central Lancashire Research laboratory. Data was analysed by Camilo L.M. Morais and Kassio M.G. Lima. Ishaan Maitra, Camilo L.M. Morais and Kassio M.G. Lima prepared a paper for publication of the results with the support of Professor Francis L. Martin.

Signed

.....
Professor F L Martin

.....
Mr I Maitra

Establishing spectrochemical changes in the natural history of oesophageal adenocarcinoma from tissue Raman mapping analysis

Introduction

Raman spectroscopy linked with variable selected methods for classification could be the potential to segregate stages of oesophageal adenocarcinoma. If dysplasia can initially be diagnosed accurately with adjuncts to histology, this would benefit earlier treatment and prevent the burden of patients developing OAC.

Raman mapping provides chemical information coupled with spatial information (Kann *et al.* 2015). Raman mapping is a non-invasive, label-free technique, with high chemical specificity that does not damage cells. Raman spectroscopy combined with microscopy is an ideal instrument for imaging biological samples and tissues.

Herein, we report three cases of OAC where initial previous pathologies of gastro-oesophageal junction (GOJ) mucosa were normal squamous epithelium in one case, and intestinal metaplasia being the initial pathology in the other two cases. The first case of OAC was diagnosed 3 months after an initially normal OGD. The second was 2 years after their previous OGD for surveillance for Barrett's oesophagus. The third was 2 ½ years after their previous OGD for surveillance for Barrett's oesophagus. There is currently no clear recognition of tissue spectrochemical markers that can distinguish between the different stages of disease in an individual patient's disease progression to oesophageal adenocarcinoma. Our aim is to understand and identify spectral differences using Raman spectroscopic mapping between both histological grades in these three illustrated cases.

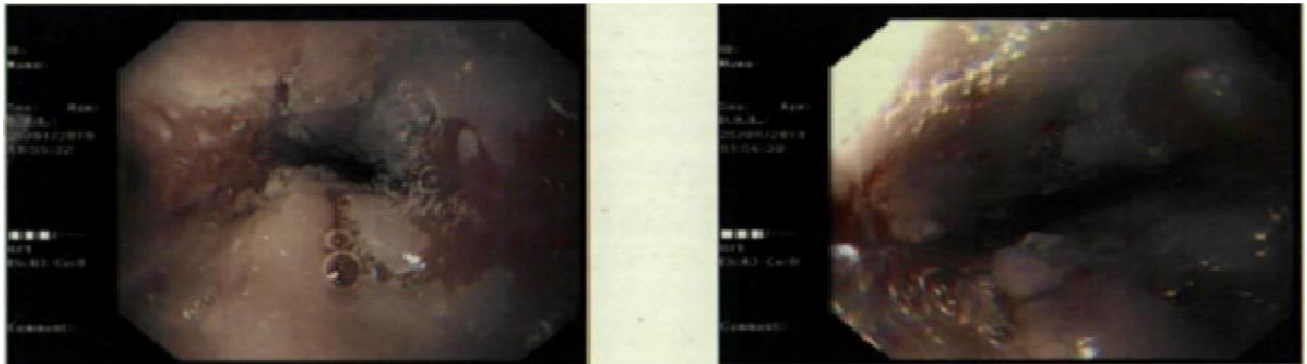
Case One (*de-novo* OAC). A 65-year-old lady presented to an upper Gastrointestinal (GI) clinic with a long-standing history of volume reflux. She had experienced dysphagia to solids and liquids over the past 2 months, with a sensation of food getting stuck at the level of her epigastrium. At the time she denied any sinister features of malignancy such as weight loss or anaemia. Her past medical history included mild Chronic Obstructive Pulmonary Disease (COPD) and hypertension. She was a non-smoker and was tee-total. An urgent upper GI endoscopy revealed mild distal oesophagitis with normal squamous epithelium confirmed on biopsy. Her Campylobacter-like organism (CLO) test was negative and she was subsequently discharged.

She presented 3 months later with a history of progressive dysphagia and a weight loss of 2 stone. Clinical examination was unremarkable and no sinister signs of pathology was seen on haematological and biochemical testing. An upper GI endoscopy however revealed a mid to distal oesophageal stricture suspicious of OAC. This was confirmed on biopsy. Further staging Computerised Tomography (CT) imaging revealed metastatic OAC (T4N2M1) with distal spread to her thoraco-lumbar spine and proximal femurs bilaterally. She had a metallic stent inserted under radiological guidance for symptomatic control. She declined further oncological input and unfortunately passed away 2 months since her malignant diagnosis.

Case Two (Barrett's oesophagus to OAC). A 69-year-old male with a background of Barrett's oesophagus and lower limb peripheral vascular disease presented at the endoscopy department for his 2-year Barrett's surveillance gastroscopy. He developed a short 1-month history of dysphagia to solids and a weight loss of 1 stone with reduced appetite prior to his surveillance gastroscopy. He was an ex-smoker (stopped 10 year ago) and was tee-total. Clinical examination and prior haematological and biochemical tests were unremarkable. His previous OGD revealed uncomplicated junctional intestinal metaplasia consistent with Barrett's oesophagus with no dysplasia. His most recent endoscopy identified a junctional OAC.

Staging CT imaging identified T3N2 disease with no evidence of distal metastatic disease. This was confirmed with staging laparoscopy performed a month after his initial endoscopic malignant diagnosis. A metallic stent was inserted under radiological guidance for symptomatic control and he has had 2 cycles of neoadjuvant chemotherapy with a view to cardio-oesophagectomy followed by adjuvant chemotherapy.

Figure 28: Ulcerative lesion on background of Barrett's oesophagus (35 cm from incisors, 1 cm length).



Case Three (Barrett's oesophagus to OAC). A 75-year-old male with a background of Barrett's oesophagus and Type II diabetes mellitus with peripheral neuropathy presented at the endoscopy department for his 2-year Barrett's surveillance gastroscopy. His past surgical history included recurrent hiatus hernia repairs resulting in exertional dyspnoea. He had no new upper GI symptoms prior to his surveillance gastroscopy. He was an ex-smoker (stopped 20 year ago) and was tee-total. Clinical examination and prior haematological and biochemical tests were unremarkable. His previous upper GI endoscopy revealed uncomplicated junctional intestinal metaplasia consistent with Barrett's oesophagus with no dysplasia. His most recent endoscopy identified findings consistent with a junctional OAC.

Staging CT imaging identified T3N1 disease with no evidence of distal metastatic disease. This was confirmed with staging laparoscopy performed two months after his initial endoscopic malignant diagnosis. The patient is currently undergoing his first cycles of neoadjuvant chemotherapy with a view to cardio-oesophagectomy followed by adjuvant chemotherapy.

Material and Methods

Ethical Approval

Ethical approval was granted by the East of England - Cambridge Central Research Ethics Committee from 2015 (Archival gastro-intestinal tissue, blood, saliva and urine collection; REC reference: 18/EE/0069; IRAS project ID: 242639) and the University of Central Lancashire (STEMH 909 application).

Pre-sampling preparation

Archival oesophageal tissue samples were acquired from the pathology laboratory at the Royal Preston Hospital from February to June 2019. Contiguous sections of 4 μ m thickness were prepared on FisherBrand™ slides (Aluminium foil covered). Contiguous samples are used so that each section closely resembles the other sections, thus ensuring correlation between histology and spectroscopic measurements. A separate Consultant Histopathologist identified sections of the cut biopsies for an overall accurate representative analytical study of the tissue. This was to ensure that spectral measurements would be taken from the appropriate area, and from the same area for the differing technologies to avoid heterogeneity in the cut tissue samples.

For each sample, the H&E section was scanned onto computer software. This allowed the regions that best reflect the overall diagnosis to be highlighted and labelled. Ten regions were selected for each sample and for each modality of analysis. Tissue blocks are embedded in paraffin wax at 20°C as this ensures durability for long-term storage without deterioration of the tissue sample. Paraffin embedding has been shown to affect 1465 cm⁻¹, 2850 cm⁻¹, 2918 cm⁻¹ and 2956 cm⁻¹ bands (Wiens *et al.*, 2007). Furthermore, paraffin wax produces a significant Raman signal with distinctive peaks (at 888cm⁻¹, 1063cm⁻¹, 1133cm⁻¹, 1171cm⁻¹, 1296cm⁻¹, 1419cm⁻¹, 1441cm⁻¹ and 1462cm⁻¹) associated with aromatic structure. These peaks are in the fingerprint region and would render the interpretation of tissue spectra impossible (Mian *et al.*, 2014).

Presently in the field of Raman spectroscopy, there is lack of consensus with regard to a standard protocol for de-paraffinization of paraffin-embedded sections (Lyng *et al.*, 2011). De-paraffinisation was hence performed prior to commencing Raman measurements using local hospital protocols employing xylene and ethanol. Once prepared the slides were transported in wooden slide boxes to the Biomedical Research Laboratory. All of the samples were stored in a de-humidified glass container to prevent condensation and physical damage.

Raman mapping measurement

Raman spectra and mapping were collected with an InVia Renishaw Raman spectrometer coupled with a charge-coupled device (CCD) detector and a Leica microscope. A 200-mW laser diode was used at a wavelength of 785 nm with a grating of 1200 lines/mm.

A silicon wafer was used to calibrate the Raman shift wavenumber value as it has a single sharp peak at 520.4 cm^{-1} , which was used as the reference point. Streamline mapping was performed by moving the sample on the motorised stage under the laser beam. The size and area of the section to be mapped was based on the regions selected by the independent Consultant Histopathologist after high-resolution H&E stain analysis. On average, ten regions from each sample were mapped with a diverse range of size area depending on the size of the area of interest. The larger maps were typically from samples of OAC, which had a single larger section of interest as compared to the other pathologies that had multiple smaller areas of interest. The measurements were made using a 785 nm laser (10% power, 30 mW) with 50× zoom magnification. For each pixel in the Raman mapping image, a Raman spectrum in the range between 725 cm^{-1} – 1813 cm^{-1} (1 cm^{-1} spectral resolution) was recorded. Average spectral point measurement tissue areas were between 2mm x 3mm and 4mm x 5mm.

Data pre-processing and analysis

The data analysis was performed within MATLAB® R2014b (MathWorks, Inc., USA) using the Classification Toolbox for MATLAB, the HYPER-Tools toolbox for MATLAB, and in-house-developed algorithms. Firstly, the three-dimensional (3D) Raman mapping images were uploaded into MATLAB and unfolded into two-dimensional (2D) structures containing n rows (number of spectra) and m columns (number of wavenumbers). Thereafter, each spectral row and column underwent pre-processing by Savitzky-Golay smoothing (21 points window, 2nd order polynomial fitting) and automatic weighted least squares (AWLS) baseline correction. The resulting pre-processed spectra were split into training (70%) and validation (30%) sets using the Kennard-Stone (1969) algorithm. These spectra were used for exploratory analysis through principal component analysis (PCA) and classification through principal component analysis linear discriminant analysis (PCA-LDA). LDA is a common classification technique used in chemometrics for Raman tissue mapping analysis (Morais *et al.*, 2019). The training set was used for model construction and optimisation. The validation set was used for final model evaluation.

PCA reduces the pre-processed spectral data to a small number of principal components (PCs) responsible for the majority of the spectral data variance (Bro and Smilde, 2014). Each PC is orthogonal to each other and they are generated in a decreasing order of explained variance, where the first PC covers most of the data variance, followed by the second PC and so on. Each PC is composed of scores and loadings, where the scores represent the variance on sample direction, hence, being used to identify similarities/dissimilarities between samples; and the loadings represent the variance on wavenumber direction, being used to identify possible spectral biomarkers responsible for class differentiation. In PCA-LDA, a LDA classifier is employed in the PCA scores space in order to systematically distinguish the samples using a Mahalanobis distance calculation (Morais and Lima, 2018). PCA-LDA models were optimised using cross-validation venetian-blinds with 10 data splits.

The PCA-LDA models output in the validation set (blind spectra) are used to calculate quality metrics or figures of merit in order to evaluate the model classification performance. Metrics such as accuracy (total number of samples correctly classified considering true and false negatives), sensitivity (proportion of positive observations correctly classified) and specificity (proportion of negative observations correctly classified) are calculated as follows (Morais and Lima, 2017):

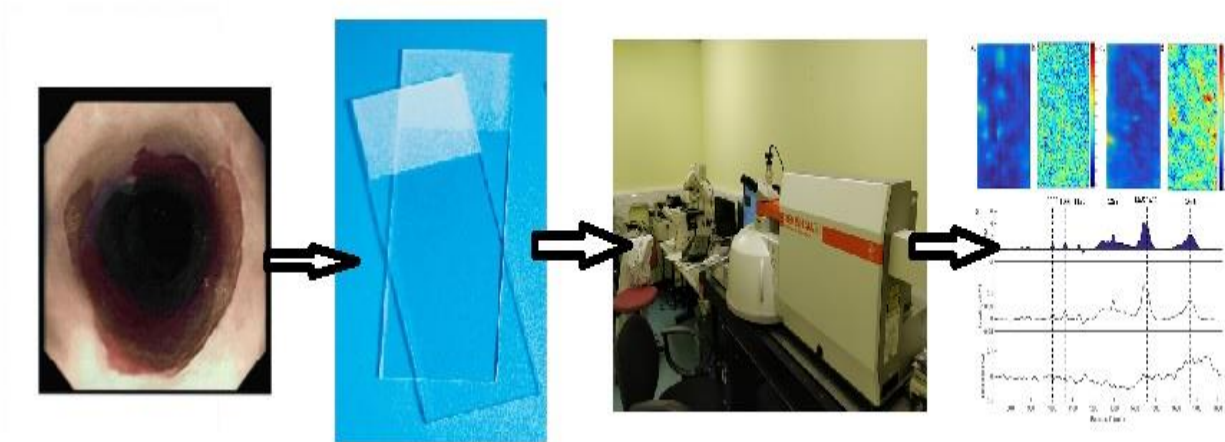
$$\text{Accuracy (\%)} = \left(\frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \right) \times 100 \quad (1)$$

$$\text{Sensitivity (\%)} = \left(\frac{\text{TP}}{\text{TP} + \text{FN}} \right) \times 100 \quad (2)$$

$$\text{Specificity (\%)} = \left(\frac{\text{TN}}{\text{TN} + \text{FP}} \right) \times 100 \quad (3)$$

where TP stands for true positives; TN for true negatives; FP for false positives; and FN for false negatives.

Figure 29: Graphical abstract demonstrating how oesophageal tissue is processed through Raman spectroscopy and mapped in order to detect oesophageal transformation stages to adenocarcinoma



Results

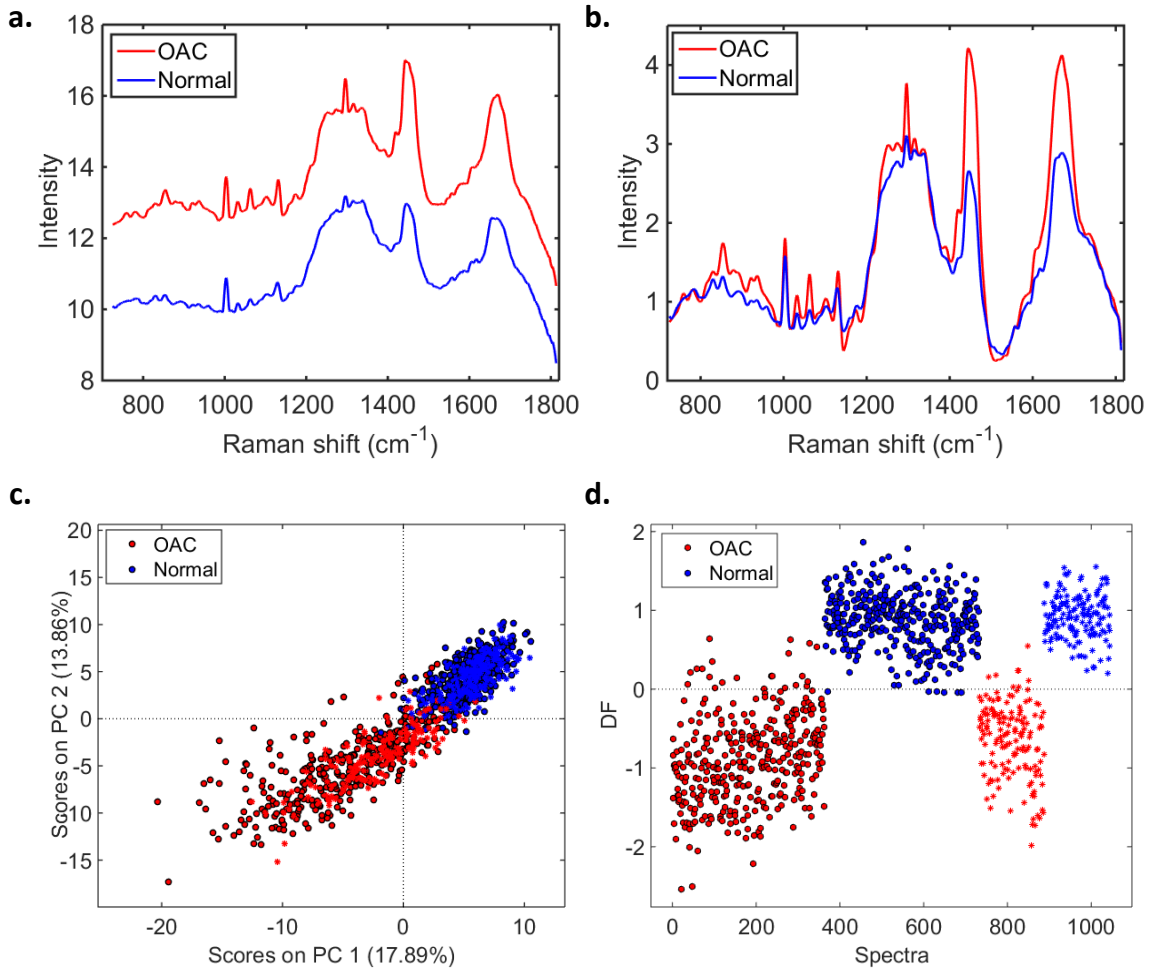
Case One (de-novo OAC)

The average raw and pre-processed (Savitzky-Golay smoothing and AWLS baseline correction) Raman spectra for normal and OAC tissue are depicted in Figure 30a and 30b, respectively. In Figure 30b there are clear spectral differences between normal and OAC tissue, especially in the regions between $800\text{ cm}^{-1} - 1000\text{ cm}^{-1}$ and $1240\text{ cm}^{-1} - 1280\text{ cm}^{-1}$, and at the peaks at 1296 cm^{-1} , 1442 cm^{-1} and 1670 cm^{-1} , where OAC has a higher intensity than the normal tissue. The PCA scores plot (Figure 30c) shows a clear natural difference between OAC and normal squamous tissue along both PC1 (17.89% explained variance) and PC2 (13.86% explained variance). A supervised classification through PCA-LDA using 5 PCs (37% explained variance) shows a very clear separation between the two tissue types (Figure 30d), where most of the spectra in the training and validation sets are correctly classified with an accuracy of 97% (94% sensitivity and 100% specificity) in validation (Table 17).

Table 17: Quality parameters (accuracy, sensitivity and specificity) for distinguishing normal vs. OAC (case 1) tissue using PCA-LDA.

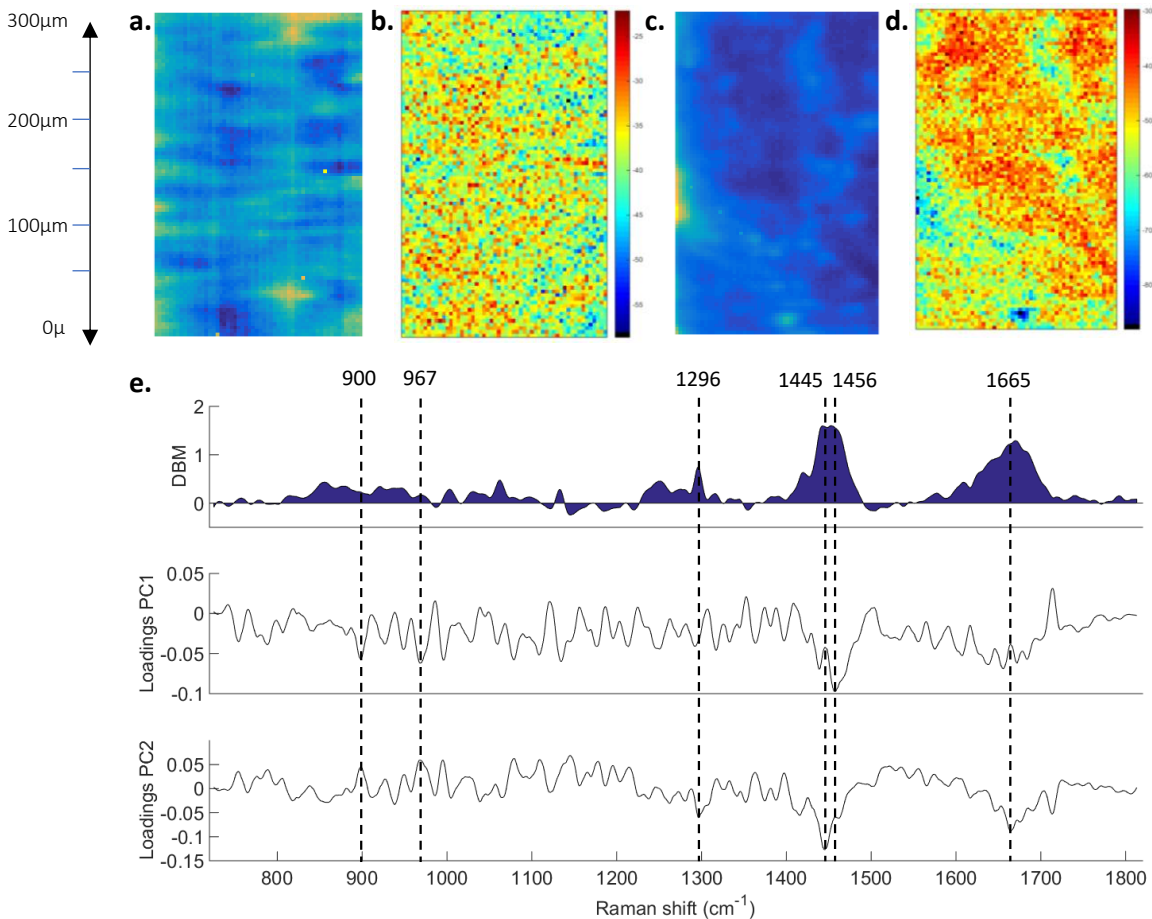
Dataset	Accuracy	Sensitivity	Specificity
Training	96%	94%	98%
Cross-validation	96%	94%	99%
Validation	97%	94%	100%

Figure 30: (a) Average raw Raman spectra for OAC and normal tissue (case 1); (b) average pre-processed (Savitzky-Golay smoothing [21 points window, 2nd order polynomial fitting] and AWLS baseline correction) Raman spectra for OAC and normal tissue (case 1); (c) PC scores plot for OAC and normal tissue; (d) PCA-LDA discriminant function (DF) plot for OAC and normal tissue (case 1), where o = training set and * = validation set of spectra.



The raw and reconstructed Raman mapping after PCA for normal tissue and OAC are shown in Figure 31a–d; and the difference-between-mean (DBM) spectrum for normal vs. OAC tissue along with the PCA loadings on PC1 and PC2 are shown in Figure 31e. The reconstructed mapping after PCA clearly shows the areas with cancerous tissue in red. Six spectral markers were found as the most important discriminant features between normal tissue and OAC: 900 cm^{-1} (C-O-C skeletal mode in monosaccharides (β -glucose)), 967 cm^{-1} (lipids), 1296 cm^{-1} (phosphodioxy (PO_2^-)), 1445 cm^{-1} (CH_2/CH_3 angular deformation in collagen), 1456 cm^{-1} (CH_2 deoxyribose) and 1665 cm^{-1} (Amide I of collagen – Movasaghi *et al.*, 2007) (Figure 31e). Peaks at around 900 cm^{-1} , 1440 cm^{-1} and 1660 cm^{-1} are indicators of cancerous tissue, as well as changes in lipids, collagen and Amide I peaks (Auner *et al.*, 2018). Changes in deoxyribose-phosphate spectral signatures have been also detected in cancer cells, which suggest partial destruction of the phosphate backbone.

Figure 31: (a) Raw and (b) PCA-recovered images for normal tissue; (c) raw and (d) PCA-recovered images for OAC tissue; (e) difference-between-mean (DBM) spectrum and PC loadings between normal tissue vs. OAC (case 1). Colour bar: mean relative intensity.



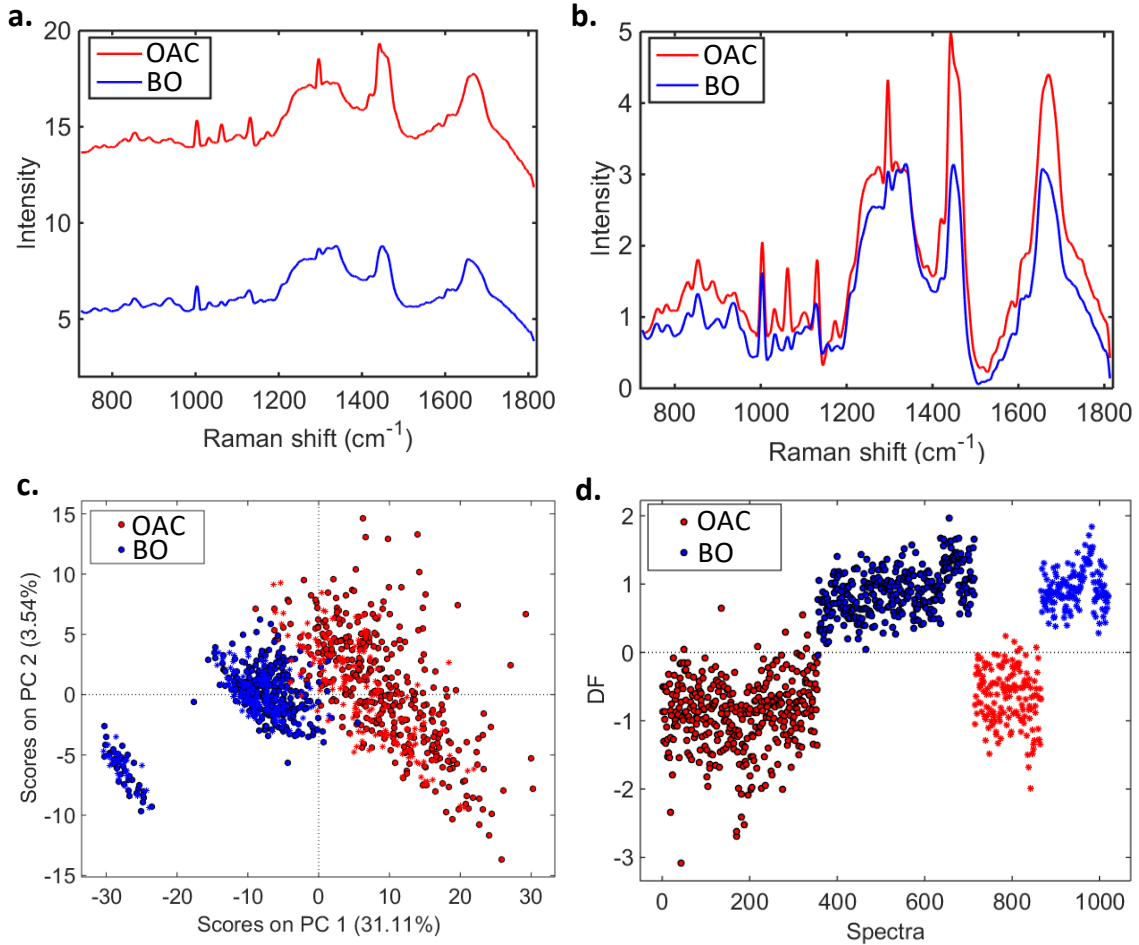
Case Two (Barrett’s oesophagus to OAC)

Figure 32a and 32b show, respectively, the average raw and pre-processed (Savitzky-Golay smoothing and AWLS baseline correction) Raman spectra for Barrett’s oesophagus and OAC tissue. In Figure 32b there are clear spectral differences between Barrett’s oesophagus and OAC tissue, where OAC tissue has an overall higher Raman intensity than Barrett’s oesophagus tissue through the whole spectrum. The PCA scores plot (Figure 32c) shows a clear natural difference between Barrett’s oesophagus and OAC tissue especially along PC1 (31.11% explained variance). PCA-LDA using 3 PCs (38% explained variance) shows a clear separation between the two tissue types (Figure 32d), where only a few OAC spectra are inside the Barrett’s oesophagus class space. This PCA-LDA model generated an accuracy of 98% (97% sensitivity and 100% specificity) to distinguish Barrett’s oesophagus tissue *vs.* OAC (Table 18).

Table 18: Quality parameters (accuracy, sensitivity and specificity) for distinguishing Barrett’s oesophagus *vs.* OAC (case 2) tissue using PCA-LDA.

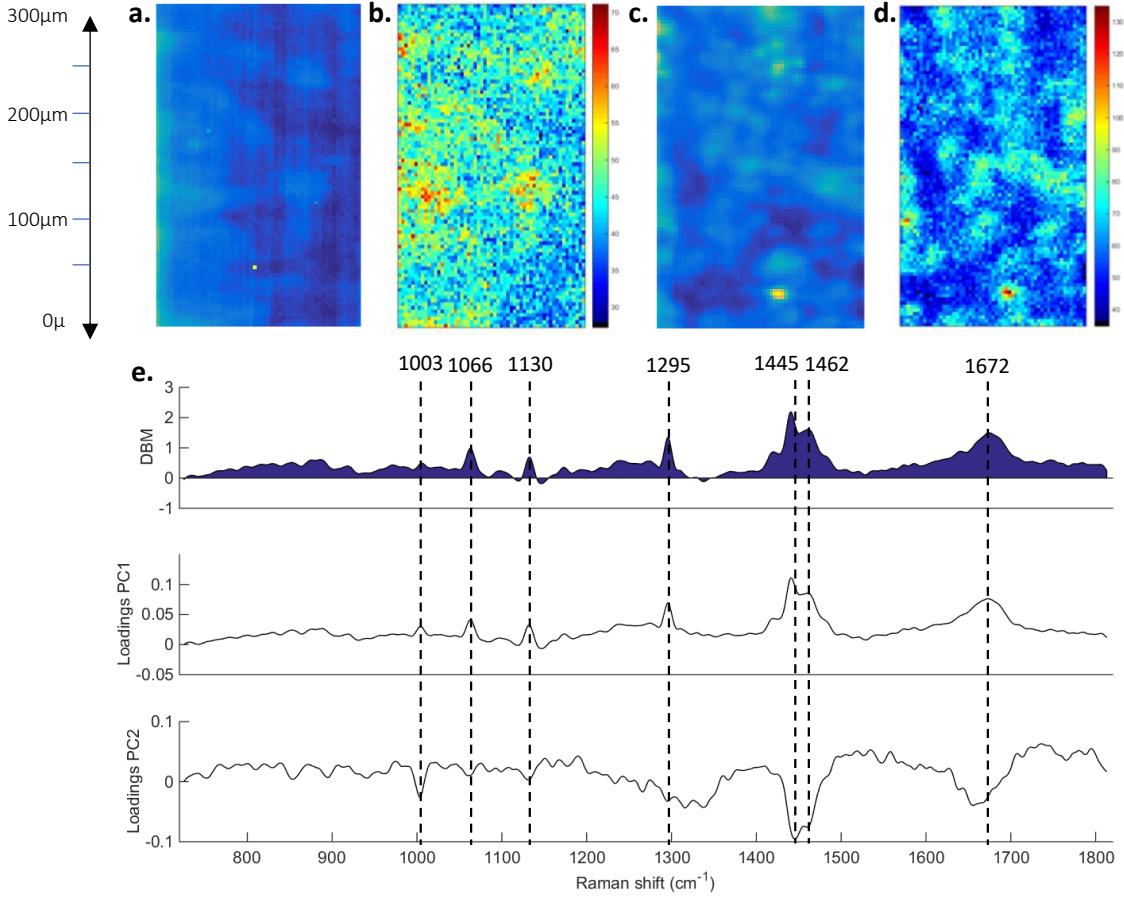
Dataset	Accuracy	Sensitivity	Specificity
Training	99%	98%	100%
Cross-validation	99%	98%	100%
Validation	98%	97%	100%

Figure 32: (a) Average raw Raman spectra for Barrett's oesophagus and OAC tissue (case 2); (b) average pre-processed (Savitzky-Golay smoothing [21 points window, 2nd order polynomial fitting] and AWLS baseline correction) Raman spectra for Barrett's oesophagus and OAC tissue (case 2); (c) PC scores plot for Barrett's oesophagus tissue and OAC tissue; (d) PCA-LDA discriminant function (DF) plot for Barrett's oesophagus and OAC tissue (case 2), where o = training set and * = validation set.



The raw and reconstructed Raman mapping after PCA for Barrett's oesophagus tissue and OAC tissue are shown in Figures 33a–d; and the DBM spectrum for Barrett's oesophagus tissue vs. OAC tissue along with the PCA loadings on PC1 and PC2 are shown in Figure 33e. The reconstructed mapping after PCA shows the areas with OAC tissue in light blue, and the Barrett's oesophagus tissue areas in yellow/red colour. The regions of OAC originate from Barrett's oesophagus tissue (intestinal metaplasia). Seven spectral markers were found as the most important discriminant features between OAC and Barrett's oesophagus tissue: 1003 cm^{-1} (C-C skeletal in phenylalanine), 1066 cm^{-1} (proline/collagen), 1130 cm^{-1} (phospholipid structural changes (*trans* vs. *gauche* isomerism)), 1295 cm^{-1} (CH_2 angular deformation), 1445 cm^{-1} (CH_2/CH_3 angular deformation in collagen), 1462 cm^{-1} (CH_2 angular deformation in disaccharides), and 1672 cm^{-1} (Amide I (C=O stretching coupled to a N-H bending)) (Figure 33e).

Figure 33: (a) Raw and (b) PCA-recovered images for Barrett’s oesophagus tissue; (c) raw and (d) PCA-recovered images for OAC tissue; (e) difference-between-mean (DBM) spectrum and PC loadings between Barrett’s oesophagus vs. OAC tissue (case 2). Colour bar: mean relative intensity.



Case Three (Barrett's oesophagus to OAC)

The average raw and pre-processed (Savitzky-Golay smoothing and AWLS baseline correction) Raman spectra for the second case of Barrett's oesophagus and OAC tissue are depicted in Figure 34a and 34b, respectively. In Figure 34b there are clear spectral differences between Barrett's oesophagus and OAC tissue, where OAC has higher Raman intensity especially in the regions between $1200\text{ cm}^{-1} - 1500\text{ cm}^{-1}$ and $1600\text{ cm}^{-1} - 1700\text{ cm}^{-1}$. The PCA scores plot (Figure 34c) shows a clear natural difference between Barrett's oesophagus and OAC tissue along PC1 (57.42% explained variance). A supervised classification through PCA-LDA using 2 PCs (60% explained variance) shows an almost perfect separation between the two tissue types (Figure 34d), where the spectra in the validation set were correctly classified with an accuracy of 100% (100% sensitivity and specificity) (Table 19).

Table 19: Quality parameters (accuracy, sensitivity and specificity) for distinguishing Barrett's oesophagus vs. OAC (case 3) tissue using PCA-LDA.

Dataset	Accuracy	Sensitivity	Specificity
Training	100%	99%	100%
Cross-validation	100%	99%	100%
Validation	100%	100%	100%

Figure 34: (a) Average raw Raman spectra for Barrett's oesophagus and OAC tissue (case 3); (b) average pre-processed (Savitzky-Golay smoothing [21 points window, 2nd order polynomial fitting] and AWLS baseline correction) Raman spectra for Barrett's oesophagus and OAC tissue (case 3); (c) PC scores plot for Barrett's oesophagus and OAC tissue (case 3); (d) PCA-LDA discriminant function (DF) plot for Barrett's oesophagus and OAC tissue (case 3), where o = training set and * = validation set.

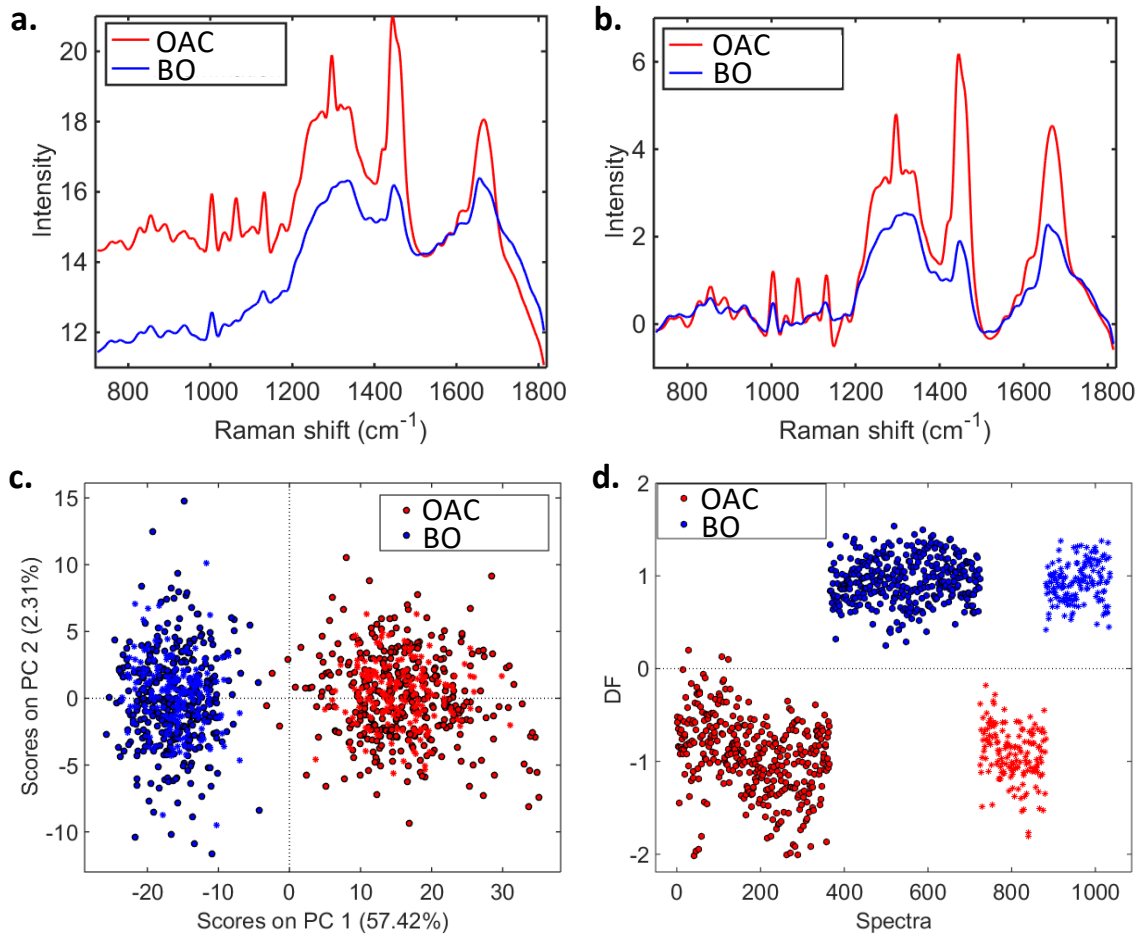
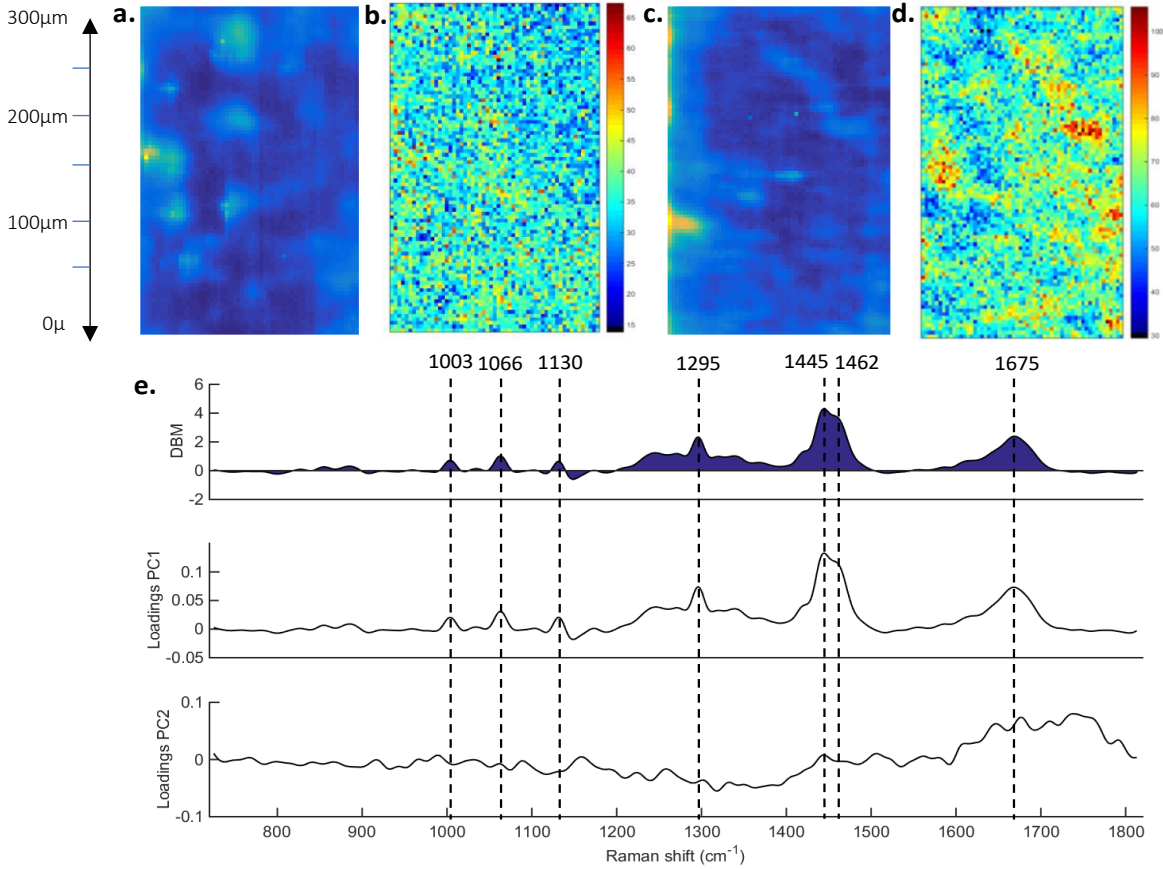


Figure 35a–d show the raw and reconstructed Raman mapping after PCA for Barrett’s oesophagus and OAC tissue (case 3); and the DBM spectrum and PCA loadings on PC1 and PC2 for Barrett’s *vs.* OAC tissue (case 3) are shown in Figure 35e. The reconstructed mapping after PCA shows the areas with OAC tissue in a higher intensity yellow/red colour, and the Barrett’s oesophagus tissue mapping in lower intensity yellow. The regions of OAC originate from Barrett’s oesophagus tissue (intestinal metaplasia). Seven spectral markers were found as the most important discriminant features between Barrett’s and OAC tissue (case 3): 1003 cm^{-1} (C-C skeletal in phenylalanine), 1066 cm^{-1} (proline/collagen), 1130 cm^{-1} (phospholipid structural changes (*trans vs. gauche* isomerism)), 1295 cm^{-1} (CH_2 angular deformation), 1445 cm^{-1} (CH_2/CH_3 angular deformation in collagen), 1462 cm^{-1} (CH_2 angular deformation in disaccharides), and 1675 cm^{-1} (Amide I). The same spectral markers observed in case 2 (Barrett’s oesophagus *vs.* OAC) were found in case 3 (Barrett’s oesophagus *vs.* OAC), confirming the consistency of this spectral methodology to provide repetitive results in different patients and that these 7 spectral markers are highly associated with a chemical difference between Barrett’s oesophagus and OAC tissue.

Figure 35: (a) Raw and (b) PCA-recovered images for Barrett's oesophagus tissue; (c) raw and (d) PCA-recovered images for OAC tissue; (e) difference-between-mean (DBM) spectrum and PC loadings between Barrett's oesophagus vs. OAC tissue (case 3). Colour bar: mean relative intensity.



Discussion

Better understanding of the carcinogenesis of Barrett's oesophagus is an essential step in targeting the disease and improving survival. The potential of present endoscopic surveillance programmes to improve detection of adenocarcinoma at early stage has been questioned by many studies (Macdonald *et al.*, 2000; Montgomery *et al.*, 2001). Case one is highly unusual as finding OAC in a patient after a normal OGD 3 months prior is rare. Furthermore, only a minority of patients progress from metaplasia to low and high-grade dysplasia (0.12-0.6% annually) (di Pietro *et al.*, 2014; Bansal and Fitzgerald, 2015). Cases two and three describe patients with an OAC diagnosis only 5 years post initial Barrett's oesophagus diagnosis. No studies to our knowledge have directly analysed spectral mapping in the same index patient, particularly in patients where the timing of diagnosis between benign disease and malignancy is ≤ 3 months. In fact, we are of the opinion that this is quite unique.

Raman spectra can be extrapolated as a direct function of the molecular composition of the tissue. Thus, there is potential that Raman can be utilised as pathological tool in validating diagnoses. There have been numerous applications of Raman spectroscopy for quantitative *ex vivo* sample analysis. These studies all range from diagnostic accuracies for establishing non-dysplastic to dysplastic tissue from 88 – 97% (Kendall *et al.*, 2003; Stone *et al.*, 2004; Chen *et al.*, 2013; Hutchings *et al.*, 2010; Almond *et al.*, 2014). This compares with our group who has established an accuracy of 97% (94% sensitivity and 100% specificity) between normal squamous epithelium and OAC. Furthermore, our group has demonstrated a 98 - 100% accuracy (98% sensitivity; 100% specificity) between Barrett's oesophagus and OAC.

Multiple studies have also qualified that the concentration of particular biomolecules elicited from Raman spectroscopy including phospholipids, proteins and collagen increases from normal squamous epithelial tissue through to dysplastic tissue (Chen *et al.*, 2013; Hutchings *et al.*, 2010; Almond *et al.*, 2014). This is in keeping with our findings as these cases have clearly demonstrated discriminant spectral markers mainly β -glucose, lipids, phosphodioxy group, deoxyribose and collagen changes associated with differences between normal squamous epithelium and OAC tissue; and phenylalanine, proline/collagen, phospholipids, disaccharides and proteins peaks associated with differences between Barrett's oesophagus and OAC tissue. The findings are particularly interesting as mapping analysis was performed directly comparing tissue in the same index patients. The presence of these discriminant subtle spectra in normal squamous epithelial tissue and intestinal metaplasia may suggest its probability in developing OAC later down the line.

Conclusion

Establishing biomolecules and biomarkers in tissue at a non-dysplastic stage can give clues to the probability of developing OAC in the future. Finding these markers early would prevent costly further invasive management requiring extensive treatment including chemoradiotherapy and/or surgical resection. This unique study has demonstrated spectrochemical differences between progressive stages to OAC on three index patients. This reinforces the potential of using Raman microspectroscopy in clinical translation, where sample diagnosis can be obtained in a computer-automated, minimally-destructive, fast and accurate manner. These preliminary results need further substantive prospective studies *in-vivo* to confirm the results and to study more biochemical components, which may be elevated with the level of dysplasia encountered.

Chapter

8

8. Discussion

Specialised intestinal metaplasia resulting from GORD is a risk factor for progression to OAC. Many patients with Barrett's oesophagus however will never progress to OAC. A meta-analysis from 2012 reported an annual incidence of OAC developing from Barrett's oesophagus of 0.33% (95% CI 0.28–0.38%) (Desai *et al.*, 2012).

At an early stage, LGD and HGD can be treated by minimally invasive techniques such as ablative therapy with a negligible risk (Phoa *et al.*, 2014). However, at an advanced stage, OAC requires invasive treatment with considerable physical burden, financial cost, and mortality (Kuipers *et al.*, 2018). Early detection and prevention are the key strategies in managing OAC. The argument as to which Barrett's oesophagus patients are most likely to benefit from surveillance and management centres on the high prevalence of Barrett's oesophagus and the low cancer incidence amongst unselected Barrett's oesophagus cases. This needs to be weighed up against the burden of invasive treatment and the high mortality associated with OAC (Kuipers *et al.*, 2018).

The 'Gold Standard' in Barrett's oesophagus identification, screening and surveillance to progression to OAC is by OGD and histopathological diagnosis. Establishing dysplasia in Barrett's oesophagus from tissue biopsy is currently the superior method in predicting development to adenocarcinoma. Possible problems with tissue biopsy analysis includes that progressive histopathological changes are subtle hence resulting in large intra and inter-observer variation in the diagnosis of dysplasia in Barrett's oesophagus. There are no defined cut-off points that distinguish disease progression when comparing inflammation, LGD and HGD. Furthermore, sampling errors can occur with small dysplasia sizes as well as its patchy distribution making sampling difficult.

Kerkhof *et al.* (2007) established that general histopathologists were found to over diagnose HGD. Nearly 40% of patients who were initially diagnosed with HGD by a general pathologist were downgraded (11% no dysplasia, 12% indefinite for dysplasia, 16% LGD) when the samples were reviewed by three experienced gastrointestinal pathologists. These results emphasise the need to obtain a second specialist GI histopathologist opinion in problematic cases. This may add delay in diagnosis increasing patient anxiety.

No clear data is available supporting the use of markers which can sub-select those at a higher risk of progression other than an expert diagnosis of LGD. Establishing a diagnosis of LGD is difficult as histopathological changes are subtle. Duits *et al.* (2015) established the overall risk of progression of LGD to HGD at 9%. Finding a distinctive biomolecular change at this stage predisposing to OAC could be utilised in risk stratification. It should be noted that even in a state of LGD, many patients may not progress to HGD and OAC.

Very few potential diagnostic and prognostic biomarkers have been shown to be reproducible and robust in the field of Barrett's oesophagus (Timmer *et al.*, 2013). Multiple ongoing studies into establishing biomarkers reflects the fact that Barrett's oesophagus needs a clinically validated prognostic tool to aid in defining risk (Bhardwaj *et al.*, 2012).

Screening and surveillance for Barrett's oesophagus involves multiple lifelong endoscopies harbouring patient anxiety with multiple invasive procedures. Furthermore, the cost-effectiveness of surveillance is often questioned because the rate of conversion from Barrett's oesophagus to adenocarcinoma is low. Biopsies taken at OGD require time and preparation for accurate diagnoses by histopathologists adding to their workload. If dysplasia could be established from biofluids instead of tissue biopsy, this could be an important milestone with regards to risk stratification and future resource planning.

Vibrational spectroscopic techniques have been used to delineate classification in oesophageal tissue from Barrett's oesophagus through to OAC in both *ex* and *in-vivo* settings utilising sophisticated spectral analysis. The majority of studies on tissue including ATR-FTIR and Raman are *ex vivo*. *In-vivo* studies using Raman probes (Bergholt *et al.*, 2011; Bergholt *et al.*, 2014) have demonstrated high specificity (>90%) in identifying dysplasia and OAC. Further high-volume multicentre studies would validate these encouraging findings.

ATR-FTIR in oesophageal tissue biopsies coupled with the SPA-QDA model was able to discriminate between normal squamous samples from 'abnormal' samples (any stage of Barrett's oesophagus) with > 90% accuracy. This multivariate classification technique (SPA-QDA) also showed promise in defining LGD in oesophageal tissue with all figures of merit for this class > 90.9%.

Raman coupled multivariate classification techniques (PCA-QDA) in identifying oesophageal stages of disease to adenocarcinoma using tissue biopsies achieved 100% sensitivity and specificity characterising LGD. These findings encourage the use of the vibrational spectroscopy as an adjunct in histopathological diagnosis in oesophageal tissue.

Our mapping case study using Raman microspectroscopy has clearly demonstrated discriminant spectral markers mainly β -glucose, lipids, phosphodioxy group, deoxyribose and collagen changes associated with differences between normal squamous epithelium and OAC tissue; and phenylalanine, proline/collagen, phospholipids, disaccharides and protein peaks associated with differences between Barrett's oesophagus and OAC tissue. Although these studies are small in volume, the presence of these subtle discriminant spectra in normal squamous epithelial tissue and intestinal metaplasia may suggest its probability into developing into OAC in the future.

Biofluid analysis has never been performed to categorise oesophageal disease states to OAC using ATR-FTIR or Raman spectroscopy. Biofluid sampling would be less invasive and less costly when compared to lifelong OGD's and biopsies.

For plasma and urine samples, the resulting GA-QDA and PCA-QDA models using ATR-FTIR successfully detected biochemical alterations as 100% for different figures of merit (accuracy, sensitivity, and F-scores) for all classifications of oesophageal disease to OAC. The method makes it possible to detect all oesophageal stages to adenocarcinoma without special sample preparation and reagents, from a minimal sample volume and shortly after sample collection.

Raman spectroscopy coupled with multivariate classification algorithms (PCA-QDA, SPA-QDA and GA-QDA) resulted in a powerful alternative approach for the detection of oesophageal stages of disease to OAC in biofluids. Saliva and urine samples were able to categorise disease processes to OAC with an excellent degree of accuracy, specificity and sensitivity (100%).

Using biofluids and spectroscopic techniques to classify oesophageal disease to OAC has a better chance of clinical adoption as, if effective, they offer considerable savings in time, money and resources.

All studies performed have used standard operating protocols with biofluid and tissue sample preparation as well as ATR-FTIR and Raman spectroscopy measurements (Butler *et al.*, 2016; Baker *et al.*, 2014). Furthermore, standardised chemometric evaluation and analysis techniques for predictive categorisation have also been utilised (Morais *et al.*, 2019). The major limitation of this thesis is the relatively small number of biofluids and tissues analysed. This would need further multicentre, multi-laboratory analysis for validation (repeatability and reproducibility of spectral datasets). Another limitation is the lack of specificity in plasma, serum, urinary and salivary ATR-FTIR and Raman spectral biomarkers in oesophageal disease processes to OAC. Other malignancies may also exhibit similar spectral changes. Further studies are also necessary to evaluate the diagnostic performance of these spectral biomarkers in other cancers.

8.1 Further work

Both ATR-FTIR and Raman spectroscopy has proven to be useful when analysing oesophageal tissue in an *ex vivo* setting. Studies using both modalities undertaken in this thesis on *ex vivo* samples endorse previous literature in the field that vibrational spectroscopy is an excellent tool in validating dysplasia and the degrees of dysplasia in Barrett's oesophagus.

The next logical step as discussed by Bergholt *et al.* (2014) would be *in vivo* Raman probes being utilised at surveillance endoscopy, providing additional targeted sampling of oesophageal mucosa. This technique would subsequently enable less tissue being sampled and reducing intubation time, thus reducing cost, patient anxiety and pathology department resources.

Our novel work on spectroscopy on biofluids has established encouraging results with both ATR-FTIR and Raman spectroscopy in plasma, saliva and urine being able to categorise classifications to OAC with a high degree of accuracy. Our studies have shown that this can potentially be utilised as a diagnostic tool for dysplasia. Using biofluids has the obvious advantage of reduced cost and being less invasive than OGD. This needs further substantial high-volume interrogation in multiple centres for validation. Translation into clinical practice relies on assay reproducibility and reliability in large sample sizes as well as a thorough cost effectiveness analysis using spectroscopy compared to endoscopic surveillance. Furthermore, the instrumentation would need to be simplified if it is to be useful in clinical settings. Subsequent biofluid studies and *in-vivo* tissue studies would validate these encouraging results thus being able to provide less invasive, less debilitating treatment associated with OAC.

9. References

- Abela, J.E., Going, J.J., Mackenzie, J.F., McKernan, M., O'Mahoney, S., Stuart, R.C. 2008, Systematic four-quadrant biopsy detects Barrett's dysplasia in more patients than nonsystematic biopsy. *Am J Gastroenterol.*, 103, pp850-855.
- Abramczyk, H., Imiela, A. 2018, The biochemical, nanomechanical and chemometric signatures of brain cancer. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 188, 8.
- Agarwal, A., Polineni, R., Hussein, Z., Vigoda, I., Bhagat, T.D., Bhattacharyya, S., Maitra, A., Verma, A. 2012, Role of epigenetic alterations in the pathogenesis of Barrett's esophagus and esophageal adenocarcinoma. *Int J Clin Exp Pathol.*, 5(5), pp382-96.
- Alexander, J.A., Jones, S.M., Smith, C.J. 1997, Usefulness of cytopathology and histology in the evaluation of Barrett's esophagus in a community hospital. *Gastrointest Endosc.*, 46, pp318–320.
- Almond, L.M., Hutchings, J., Shepherd, N., Barr, H., Stone, N., Kendall, C. 2011, Raman spectroscopy: a potential tool for early objective diagnosis of neoplasia in the oesophagus. *J Biophotonics.*, 4(10), pp685-695.
- Almond, L.M., Hutchings, J., Lloyd, G., Barr, H., Shepherd, N., Day, J., Stevens, O., Sanders, S., Wadley, M., Stone, N., Kendall, C. 2014, Endoscopic Raman spectroscopy enables objective diagnosis of dysplasia in Barrett's esophagus. *Gastrointest Endosc.*, 79(1), pp37-45.
- Amadi, C., Gatenby, P. 2017, Barrett's oesophagus: Current controversies. *World J Gastroenterol.*, 23(28), pp5051–5067.
- Andleed, F., Janjua, H.U., Atiq, A., Atiq, M., Malik, S. 2018, Attenuated total reflection spectroscopy to diagnose skin cancer and to distinguish different metastatic potential of melanoma cell. *Cancer Biomark.*, 1, pp1–8.
- Aroca, R.F. 2004, Surface-enhanced infrared spectroscopy. *Appl. Spectrosc.*, 58(11), pp324A–338A.
- Atherfold, P.A., Jankowski, J.A. 2006, Molecular biology of Barrett's cancer. *Best Pract Res Clin Gastroenterology.*, 20, pp813–827.
- Auner, G.W., Koya, S.K., Huang, C., Broadbent, B., Trexler, M., Auner, Z., Elias, A., Mehne, K.C., Brusatori, M.A. 2018, Applications of Raman spectroscopy in cancer diagnosis. *Cancer Metastasis Rev.*, 37(4), pp691-717.
- Baker, M.J., Trevisan, J., Bassan, P., Bhargava, R., Butler, H.J., Dorling, K.M., Fielden, P.R., Fogarty, S.W., Fullwood, N.J., Heys, K.A., Hughes, C., Lasch, P., Martin-Hirsch, P.L., Obinaju, B., Sockalingum, G.D., Sulé-Suso, J., Strong, R.J., Walsh, M.J., Wood, B.R., Gardner, P., Martin, F.L. 2014, Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc.*, 9(8), pp1771-91.
- Bansal, A., Fitzgerald, R.C. 2015, Biomarkers in Barrett's Esophagus: Role in Diagnosis, Risk Stratification, and Prediction of Response to Therapy. *Gastroenterology clinics of North America.*, 44, pp373–390.

- Berger, A.J., Koo, T.W., Itzkan, I., Horowitz, G., Feld, M.S. 1999, Multicomponent blood analysis by near-infrared Raman spectroscopy. *Appl Opt.*, 38, pp2916–2926
- Bergholt, M.S., Zheng, W., Lin, K., Ho, K.Y., The, M., Yeoh, K.G., So, J.B., Huang, Z. 2011, In vivo diagnosis of esophageal cancer using image-guided Raman endoscopy and biomolecular modeling. *Technol Cancer Res Treat*, 10(2), pp103-12.
- Bergholt, M.S., Zheng, W., Ho, K.Y., The, M., Yeoh, K.G., Yan So, J.B., Shabbir, A., Huang, Z. 2014, Fiberoptic confocal raman spectroscopy for real-time in vivo diagnosis of dysplasia in Barrett's esophagus. *Gastroenterology.*, 146(1), pp27–32.
- Bhardwaj, A., Hollenbeak, C.S., Pooran, N., Mathew A. 2009, A meta-analysis of the diagnostic accuracy of esophageal capsule endoscopy for Barrett's esophagus in patients with gastroesophageal reflux disease. *Am J Gastroenterol.*, 104(6), pp1533-1539.
- Bhardwaj, A., McGarrity, T.J., Stairs, D.B., Mani, H. 2012, Barrett's Esophagus: Emerging Knowledge and Management Strategies, *Pathol Res Int.*, 2012, 814146.
- Bhat, S., Coleman, H.G., Yousef, F., Johnston, B.T., McManus, D.T., Gavin, A.T., Murray, L.J. 2011, Risk of malignant progression in Barrett's Esophagus patients: results from a large population-based study. *J Natl Cancer Inst.*, 103(13), pp1049–57.
- Blot, W.J., McLaughlin, J.K. 1999. The changing epidemiology of esophageal cancer. *Semin Oncol.*, 26(5;15), pp2–8.
- Booth, C.L., Thompson, K.S. 2012, Barrett's esophagus: A review of diagnostic criteria, clinical surveillance practices and new developments. *J Gastrointest Oncol.*, 3(3), pp232–242.
- Borovicka, J., Fischer, J., Neuweiler, J., Netzer, P., Gschossmann, J., Ehmann, T., Bauerfeind, P., Dorta, G., Zürcher, U., Binek, J., Meyenberger, C. 2006, Autofluorescence endoscopy in surveillance of Barrett's esophagus: a multicenter randomized trial on diagnostic efficacy. *Endoscopy.*, 38(9), pp867-872.
- Bosher, L.H., Taylor, F.H. 1951, Heterotopic gastric mucosa in the esophagus with ulceration and stricture formation. *J Thorac Surg.*, 21, pp306–312.
- Boyer, J., Laugier, R., Chemali, M., Arpurt, J.P., Boustière, C., Canard, J.M., Dalbies, P.A., Gay, G., Escourrou, J., Napoléon, B., Palazzo, L., Ponchon, T., Richard-Mollard, B., Sautereau, D., Tucac, G., Vedrenne, B., French Society of Digestive Endoscopy SFED. 2007, French society of digestive endoscopy SFED guideline: monitoring of patients with Barrett's esophagus, *Endoscopy.*, 39(9), pp840–842.
- Bro, R., Smilde, A.K. 2014, Principal component analysis. *Anal. Methods.*, 6, pp2812-2831.
- Bus, P., Kestens, C., Ten Kate, F.J.W. 2016, Profiling of circulating microRNAs in patients with Barrett's esophagus and esophageal adenocarcinoma. *J Gastroenterol.*, 51, pp560–570.
- Butler, H. J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J., Esmonde-White, K., Fullwood, N. J., Gardner, B., Martin-Hirsch, P. L., Walsh, M. J., McAnish, M. R., Stone, N., Martin, F. L. 2016, Using Raman spectroscopy to characterize biological material. *Nature Protocols.* 11, pp. 664-687

- Buttar, N.S., Wang, K.K., Sebo, T.J., Riehle, D.M., Krishnadath, K.K., Lutzke, L.S., Anderson, M.A., Petterson, T.M., Burgart, L.J. 2001, Extent of high-grade dysplasia in Barrett's esophagus correlates with risk of adenocarcinoma. *Gastroenterology.*, 120(7), pp1630-1639.
- Cameron, A. J., Carpenter, H. A. 1997, Barrett's esophagus, high-grade dysplasia, and early adenocarcinoma: a pathological study. *Am. J. Gastroenterol.*, 92(4), pp586-591.
- Chan, J.W., Taylor, D.S., Zwerdling, T., Lane, S.M., Ihara, K., Huser, T. 2006, Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells. *Biophys. J.*, 90, pp648-656.
- Chen, K., Adato, R., Altug, H. 2012, Dual-band perfect absorber for multispectral plasmonenhanced infrared spectroscopy. *ACS Nano.*, 6(9), pp7998–8006.
- Chen, L., Wang, Y., Liu, N., Lin, D., Weng, C., Zhang, J., Zhu, L., Chen, W., Chen, R., Feng, S. 2013, Near-infrared confocal micro-Raman spectroscopy combined with PCA-LDA multivariate analysis for detection of esophageal cancer. *Laser Phys.*, 23(6).
- Cheung, K.T., Trevisan, J., Kelly, J.G., Ashton, K.M., Stringfellow, H.F., Taylor, S.E., Singh, M.N., Martin-Hirsch, P.L., Martin F.L. 2011, Fourier-transform infrared spectroscopy discriminates a spectral signature of endometriosis independent of inter-individual variation. *Analyt.*, 136, pp2047.
- Cook, M.B., Shaheen, N.J., Anderson, L.A., Giffen, C., Chow, W.H., Vaughan, T.L., Whiteman, D.C., Corley, D.A. 2012, Cigarette smoking increases risk of Barrett's esophagus: an analysis of the Barrett's and Esophageal Adenocarcinoma Consortium. *Gastroenterology.*, 142(4), pp744–753.
- Corley, D.A., Kubo, A., Levin, T.R., Block, G., Habel, L., Zhao, W., Leighton, P., Quesenberry, C., Rumore, G.J., Buffler, P.A. 2007, Abdominal obesity and body mass index as risk factors for Barrett's esophagus. *Gastroenterology.*, 133(1), pp34–41.
- Corley, D.A., Kubo, A., Levin, T.R., Block, G., Habel, L., Rumore, G., Quesenberry C., Buffler, P. 2009, Race, ethnicity, sex and temporal differences in Barrett's oesophagus diagnosis: a large community-based study, 1994–2006. *Gut.*, 58(2), pp182–188.
- Curvers W., Baak L., Kiesslich, R., Van Oijen, A., Rabenstein, T., Ragnath, K., Rey, J.F., Scholten, P., Seitz, U., Ten Kate, F., Fockens, P., Bergman, J. 2008, Chromoendoscopy and narrow-band imaging compared with high-resolution magnification endoscopy in Barrett's esophagus. *Gastroenterology.*, 134(3), pp670–9.
- DaCosta, R.S., Wilson, B.C., Marcon, N.E. 2003, Photodiagnostic techniques for the endoscopic detection of premalignant gastrointestinal lesions. *Digestive Endoscopy.*, 15, pp153-173.
- DaCosta, R.S., Wilson, B.C., Marcon, N.E. 2006, Spectroscopy and fluorescence in esophageal diseases. *Best Pract Res Clin Gastroenterol.*, 20, pp41-57.
- Davis, V.W, Schiller, D.E., Eurich, D., Sawyer, M.B. 2012, Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. *World J. Surg. Oncol.*, 10, pp271.
- Dawes, C., Pedersen, A.M., Villa, A., Ekström, J., Proctor, G.B., Vissink, A., Aframian, D., McGowan, R., Aliko, A., Narayana, N., Sia, Y.W., Joshi, R.K., Jensen, S.B., Kerr, A.R., Wolff, A. 2015, The functions of human saliva: A review sponsored by the World Workshop on Oral Medicine VI. *Arch Oral Biol.*, 60(6), pp863-74.

- de Jonge, P.J., Van Blankenstein, M., Looman, C.W., Casparie, M.K., Meijer, G.A., Kuipers, E.J. 2010, Risk of malignant progression in patients with Barrett's oesophagus: a Dutch nationwide cohort study. *Gut.*, 59(8), pp1030-1036.
- den Hoed, C.M., van Blankenstein, M., Dees, J., Kuipers, E.J. 2011, The minimal incubation period from the onset of Barrett's oesophagus to symptomatic adenocarcinoma. *Br J Cancer.*, 105, pp200–205.
- Desai, T.K., Krishnan, K., Samala, N., Singh, J., Cluley, J., Perla, S., Howden, C.W. 2012, The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut.*, 61, pp970-976.
- Diem, M., Ergin, A., Remiszewski, S., Mu, X., Akalin, A., Raz, D. 2016, Infrared micro-spectroscopy of human tissue: principles and future promises. *Faraday Discuss.*, 187, pp9-42.
- di Pietro, M., Alzoubaidi, D., Fitzgerald, R.C. 2014, Barrett's esophagus and cancer risk: how research advances can impact clinical practice. *Gut and liver.*, 8, pp356–370.
- Duits, L.C., Lao-Sirieix, P., Wolf, W.A., O'Donovan, M., Galeano-Dalmau, N., Meijer, S.L., Offerhaus, G.J.A., Redman, J., Crawte, J., Zeki, S., Pouw, R.E., Chak, A., Shaheen, N.J., Bergman, J.J.G.H.M., Fitzgerald, R.C. 2019, A biomarker panel predicts progression of Barrett's esophagus to esophageal adenocarcinoma. *Dis Esophagus.*, 32(1), pp1.
- Dulak, A.M., Schumacher, S.E., Van Lieshout, J., Imamura, Y., Fox, C., Shim, B., Ramos, A.H., Saksena, G., Baca, S.C., Baselga, J., Taberero, J., Barretina, J., Enzinger, P.C., Corso, G., Roviello, F., Lin, L., Bandla, S., Luketich, J.D., Pennathur, A., Meyerson, M., Ogino, S., Shivdasani, R.A., Beer, D.G., Godfrey, T.E., Beroukhim, R., Bass, A.J. 2012, Gastrointestinal adenocarcinomas of the esophagus, stomach, and colon exhibit distinct patterns of genome instability and oncogenesis. *Cancer Res.*, 72(17), pp4383–4393.
- Dulak, A.M., Stojanov, P., Peng, S., Lawrence, M.S., Fox, C., Stewart, C., Bandla, S., Imamura, Y., Schumacher, S.E., Shefler, E., McKenna, A., Carter, S.L., Cibulskis, K., Sivachenko, A., Saksena, G., Voet, D., Ramos, A.H., Auclair, D., Thompson, K., Sougnez, C., Onofrio, R.C., Guiducci, C., Beroukhim, R., Zhou, Z., Lin, L., Lin, J., Reddy, R., Chang, A., Landrenau, R., Pennathur, A., Ogino, S., Luketich, J.D., Golub, T.R., Gabriel, S.B., Lander, E.S., Beer, D.G., Godfrey, T.E., Getz, G., Bass, A.J. 2013, Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat Genet.*, 45(5), pp478-486.
- Dunbar, K.B., Spechler, S.J. 2012. The risk of lymph-node metastases in patients with high-grade dysplasia or intramucosal carcinoma in Barrett's esophagus: a systematic review. *Am J Gastroenterol.*, 107(6), pp850–862.
- Eilers, P.H.C. 2004, A perfect smoother. *Anal. Chem.*, 76, pp404–11.
- Eloubeidi, M.A., Provenzale, D. 1999. Does this patient have Barrett's esophagus? The utility of predicting Barrett's esophagus at the index endoscopy. *Am J Gastroenterol.*, 94, pp937-943.
- El-Serag, H.B., Naik, A.D., Duan, Z. 2016, Surveillance endoscopy is associated with improved outcomes of oesophageal adenocarcinoma detected in patients with Barrett's oesophagus. *Gut.*, 65, pp1252–1260.

- Evans, J.A., Poneros, J.M., Bouma, B.E., Bressner, J., Halpern, E.F., Shishkov, M., Lauwers, G.Y., Mino-Kenudson, M., Nishioka, N.S., Tearney, G.J. 2006, Optical Coherence Tomography to Identify Intramucosal Carcinoma and High-Grade Dysplasia in Barrett's Esophagus. *Clin Gastroenterol Hepatol.*, 4(1), pp38-43.
- Fadaka, A., Ajiboye, B., Ojo, O., Adewale, O., Olayide, I., Emuowhochere, R. 2017, Biology of glucose metabolism in cancer cells. *Journal of Oncology Sciences.*,3, pp45.
- Faoláin, E.O., Hunter, M.B., Byrne, J.M., Kelehan, P., Lambkin, H.A., Byrne, H.J., Lyng, F.M. 2005, Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents. *J Histochem Cytochem.* 2005, 53(1), pp121-129.
- Feng, S., Chen, R., Lin, J., Pan, J., Chen, G., Li, Y., Cheng, M., Huang, Z., Chen, J., Zeng, H. 2010, Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. *Biosens Bioelectron.*, 25, pp2414–2419.
- Fitzgerald, R.C., Onwuegbusi, B.A., Bajaj-Elliott, M., Saeed, I.T., Burnham, W.R., Farthing, M.J. 2002, Diversity in the oesophageal phenotypic response to gastro-oesophageal reflux: immunological determinants. *Gut.*, 50(4), pp451-459.
- Fitzgerald, R.C., di Pietro, M., Ragunath, K., Ang, Y., Kang, J.Y., Watson, P., Trudgill, N., Patel, P., Kaye, P.V., Sanders, S., O'Donovan, M., Bird-Lieberman, E., Bhandari, P., Jankowski, J.A., Attwood, S., Parsons, S.L., Loft, D., Lagergren, J., Moayyedi, P., Lyratzopoulos, G., de Caestecker, J. 2014, British Society of Gastroenterology. British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut.*, 63(1), pp7-42.
- Fortun, P.J., Anagnostopoulos, G.K., Kaye, P., James, M., Foley, S., Samuel, S., Shonde, A., Badreldin, R., Campbell, E., Hawkey, C.J., Ragunath, K. 2006, Acetic acid-enhanced magnification endoscopy in the diagnosis of specialized intestinal metaplasia, dysplasia and early cancer in Barrett's oesophagus. *Aliment Pharmacol Ther.*, 23(6), pp735–42.
- Fouad, Y.M., Mostafa, I., Yehia, R., El-Khayat, H. 2014, Biomarkers of Barrett's esophagus. *World J Gastrointest Pathophysiol.*, 5(4), pp450–456.
- Gajjar, K., Trevisan, J., Owens, G., Keating, P.J., Wood, N.J., Stringfellow, H.F., Martin-Hirsch, P.L., Martin, F.L. 2013, Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. *Analyst.*, 138(14), pp3917-26.
- Gindea, C., Birla, R., Hoara, P., Caragui, A., Constantinoiu, S. 2014, Barrett Esophagus: History, definition and etiopathogeny. *J Med Life.*, 7(3), pp23–30.
- González-Solís, J.L, Martínez-Espinosa, J.C., Torres-González, L.A., Aguilar-Lemarroy, A., Jave-Suárez, L.F., Palomares-Anda, P. 2014, Cervical cancer detection based on serum sample Raman spectroscopy. *Lasers Med. Sci.*, 2014, 29, 979.
- Gordon, L.G., Mayne, G.C., Hirst, N.G., Bright, T., Whiteman, D.C., Watson, D.I. 2014, Cost-effectiveness of endoscopic surveillance of non-dysplastic Barrett's esophagus. *Gastrointest Endosc.*, 79, pp242–252.

- Grady, W.M., Yu, M. 2018, Molecular Evolution of Metaplasia to Adenocarcinoma in the Esophagus. *Dig Dis Sci.*, 63(8), pp2059-2069.
- Graham, D., Faulds, K. 2009, Surface-enhanced Raman scattering as a detection technique for molecular diagnostics. *Expert. Rev. Mol. Diagn.*, 9, pp537.
- Graham, D., Lipman, G., Sehgal V., Lovat, L.B. 2016, Monitoring the premalignant potential of Barrett's oesophagus. *Frontline Gastroenterol.*, 7(4), pp316–322.
- Guder, W.G., Narayanan, S., Wisser, H., Zawata, B. 2003, Samples: from the patient to the laboratory: the impact of preanalytical variables on the quality of laboratory results. 3rd Edition. Weinheim, New York: Wiley-VCH.
- Hands, J.R., Dorling, K.M., Abel, P., Ashton, K.M., Brodbelt, A., Davis, C., Dawson, T., Jenkinson, M.D., Lea, R.W., Walker, C., Baker, M.J. 2014, Attenuated total reflection fourier transform infrared (ATR-FTIR) spectral discrimination of brain tumour severity from serum samples. *J Biophotonics.*, 7(3-4), pp189-99.
- Hanna, S., Rastogi, A., Weston, A.P., Totta, F., Schmitz, R., Mathur, S., McGregor, D., Cherian, R., Sharma, P. 2006, Detection of Barrett's esophagus after endoscopic healing of erosive esophagitis. *Am J Gastroenterol.*, 101(7), pp1416-20.
- Harrison, R., Perry, I., Haddadin, W., McDonald, S., Bryan, R., Abrams, K., Sampliner, R., Talley, N.J., Moayyedi, P., Jankowski, J.A. 2007, Detection of intestinal metaplasia in Barrett's esophagus: an observational comparator study suggests the need for a minimum of eight biopsies. *Am J Gastroenterol.*, 102(6), pp1154-1161.
- Harvey, T.J., Faria, E.C., Henderson, A., Gazi, E., Ward, A.D., Clarke, N.W., Brown, M.D., Snook, R.D., Gardner, P. 2008, Spectral discrimination of live prostate and bladder cancer cell lines using Raman optical tweezers. *J. Biomed. Opt.*, 13, pp234-345.
- Hayeck, T.J., Kong, C.Y., Spechler, S.J., Gazelle, G.S., Hur, C. 2010, The prevalence of Barrett's esophagus in the US: estimates from a simulation model confirmed by SEER data. *Dis Esophagus.* 23, pp451–457.
- He, H., Tian, D., Guo, J., Liu, M., Chen, Z., Hamdy, F.C., Helleday, T., Su, M., Ying, S. 2013, DNA damage response in peritumoral regions of oesophageal cancer microenvironment. *Carcinogenesis.*, 34(1), pp139–145.
- Heberle, C.R., Omidvari, A.H., Ali, A., Kroep, S., Kong, C.Y., Inadomi, J.M., Rubenstein, J.H., Tramontano, A.C., Dowling, E.C., Hazelton, W.D., Luebeck, E.G., Lansdorp-Vogelaar, I., Hur, C. 2017, Cost effectiveness of screening patients with gastroesophageal reflux disease for Barrett's esophagus with a minimally invasive cell sampling device. *Clin Gastroenterol Hepatol.*, 15(9), pp1397–1404.e7
- Hiremath, G.S., Powers, J., Sivakumar, A., Correa, H., Acra, S., Mahadevan-Jansen, A. Raman Microspectroscopy Reveals Distinct Biochemistry of Esophageal biopsies Obtained from Children with Eosinophilic Esophagitis. In: 4th CURED EGID Research Conference and Patient Education Program. Cincinnati; 2017
- Hirota, W.K., Loughney, T.M., Lazas, D.J., Maydonovitch, C.L., Rholl, V., Wong, R.K. 1999, Specialized intestinal metaplasia, dysplasia, and cancer of the esophagus and esophagogastric junction: prevalence and clinical data. *Gastroenterology.*, 116, pp277-285.

- Huo, X., Juergens, S., Zhang, X., Rezaei, D., Yu, C., Strauch, E.D., Wang, J.Y., Cheng, E., Meyer, F., Wang, D.H., Zhang, Q., Spechler, S.J., Souza, R.F. 2011, Deoxycholic acid causes DNA damage while inducing apoptotic resistance through NF- κ B activation in benign Barrett's epithelial cells. *Am J Physiol Gastrointest Liver Physiol.*, 301(2), ppG278-G286.
- Hutchings, J., Kendall, C., Barr, H., Stone, N., Shepherd, N. 2010, Evaluation of linear discriminant analysis for automated Raman histological mapping of esophageal high-grade dysplasia. *J. Biomed. Opt.*, 15,066015.
- Hvid-Jensen, F., Pedersen, L., Drewes, A.M., Sorensen, H.T., Funch-Jensen, P. 2011, Incidence of adenocarcinoma among patients with Barrett's esophagus. *NEJM.*, 365(15). pp1375–83.
- Inadomi, J.M., Sampliner, R., Lagergren, J., Lieberman, D., Fendrick, A.M., Vakil, N. 2003, Screening and surveillance for Barrett esophagus in high-risk groups: a cost–utility analysis. *Ann Intern Med.*, 138(3), pp176-186.
- Janmaat, V.T., van Olphen, S.H., Biermann, K.E., Looijenga, L.H.J, Bruno, M.B., Spaander, M.C.W. 2017, Use of immunohistochemical biomarkers as independent predictor of neoplastic progression in Barrett's oesophagus surveillance: A systematic review and meta-analysis. *PLoS One.*, 12(10), e0186305.
- Jobe, B.A., Hunter, J.G., Chang, E.Y. 2006, Office-based unsedated small-caliber endoscopy is equivalent to conventional sedated endoscopy in screening and surveillance for Barrett's esophagus: a randomized and blinded comparison. *Am J Gastroenterol.*, 101, pp2693–2703.
- Johansson, J., Johnsson, F., Walther, B., Willen, R., vonHolstein, C.S., Zilling, T. 1996, Adenocarcinoma in the distal esophagus with and without Barrett esophagus. Differences in symptoms and survival rates. *Arch Surg.*, 131, pp708–13.
- Kadri, S.R., Lao-Sirieix, P., O'Donovan, M., Debiram, I., Das, M., Blazeby, J.M., Emery, J., Boussioutas, A., Morris, H., Walter, F.M., Pharoah, P., Hardwick, R.H., Fitzgerald, R.C. 2010, Acceptability and accuracy of a non-endoscopic screening test for Barrett's oesophagus in primary care: cohort study. *BMJ.*, 341, c4372.
- Kann, B., Offerhaus, H.L., Windbergs, M., Otto, C. 2015, Raman microscopy for cellular investigations—from single cell imaging to drug carrier uptake visualization. *Adv Drug Deliv Rev.*, 89, pp71–90.
- Kara, M.A., Ennahachi, M., Fockens, P., ten Kate, F.J., Bergman, J.J. 2006, Detection and classification of the mucosal and vascular patterns (mucosal morphology) in Barrett's esophagus by using narrow band imaging. *Gastrointest Endosc.*, 64(2), pp155-166.
- Kara, M.A., DaCosta, R.S., Streutker, C.J., Marcon, N.E., Bergman, J.J., Wilson, B.C. 2007, Characterization of tissue autofluorescence in Barrett's esophagus by confocal fluorescence microscopy. *Dis Esophagus.*, 20(2), pp141-150.
- Kara, M.A., Peters, F.P., Rosmolen, W.D., Krishnadath, K.K., ten Kate, F.J., Fockens, P., Bergman, J.J. 2005, High-resolution endoscopy plus chromoendoscopy or narrow-band imaging in Barrett's esophagus: a prospective randomized crossover study. *Endoscopy.*, 37 (10), pp929-36. (A)

- Kara, M.A., Smits, M.E., Rosmolen, W.D., Bultje, A.C., Ten Kate, F.J., Fockens, P., Tytgat, G.N., Bergman, J.J. 2005, A randomized crossover study comparing light-induced fluorescence endoscopy with standard videoendoscopy for the detection of early neoplasia in Barrett's esophagus. *Gastrointest Endosc.*, 61(6), pp671-678. (B)
- Kastelein, F., van Olphen, S., Steyerberg, E.W., Sikkema, M., Spaander, M.C., Looman, C.W., Kuipers, E.J., Siersema, P.D., Bruno, M.J., de Bekker-Grob, E.W.; ProBar-study group. 2015, Surveillance in patients with long-segment Barrett's oesophagus: a cost-effectiveness analysis. *Gut.*, 64(6), pp864-71.
- Kaye, P.V., Haider, S.A., James, P.D., Soomro, I., Catton, J., Parsons, S.L., Ragunath, K., Ilyas, M. 2010, Novel staining pattern of *p53* in Barrett's dysplasia--the absent pattern. *Histopathology.*, 57, pp933-5.
- Kelly, J.G., Ahmadzai, A.A., Hermansen, P., Pitt, M.A., Saidan, Z., Martin-Hirsch, P.L., Martin F.L. 2011, A biospectroscopic interrogation of fine needle aspirates points towards segregation between graded categories: an initial study towards diagnostic screening. *Anal. Bioanal. Chem.*, 401, pp957-967.
- Kendall, C., Stone, N., Shepherd, N., Geboes, K., Warren, B., Bennett, R., Barr, H. 2003, Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. *J Pathol.*, 200(5), pp602-9.
- Kennard, R., Stone, L. 1969, Computer aided design of experiments. *Technometric.*, 11, pp137-148.
- Kerkhof, M., van Dekken, H., Steyerberg, E.W., Meijer, G.A., Mulder, A.H., de Bruïne, A., Driessen, A., ten Kate, F.J., Kusters, J.G., Kuipers, E.J., Siersema, P.D.; CYBAR study group. 2007, Grading of dysplasia in Barrett's oesophagus: substantial interobserver variation between general and gastrointestinal pathologists. *Histopathology.*, 50(7), pp920-927.
- Kestens C, Offerhaus GJ, van Baal JW, Siersema PD. 2015, Patients with Barrett's esophagus and persistent low-grade dysplasia have an increased risk for high-grade dysplasia and cancer. *Clin Gastroenterol Hepatol.*, 14(7):956-962.e1.
- Khan, S., Do, K.A., Kuhnert, P., Pillay, S.P., Papadimos, D., Conrad, R., Jass, J.R. 1998, Diagnostic value of *p53* immunohistochemistry in Barrett's esophagus: an endoscopic study. *Pathology.*, 30, pp136-40.
- Kneipp, K., Haka, A.S., Kneipp, H., Badizadegan, K., Yoshizawa, N., Boone, C., Feld, M.S. 2002, Surface-enhanced Raman spectroscopy in single living cells using gold nanoparticles. *Appl. Spectrosc.*, 56(2), pp150-154.
- Krishnamoorthi, R., Singh, S., Ragunathan, K., Visrodia, K., Wang, K.K., Katzka, D.A., Iyer, P.G. 2018, Factors associated with progression of Barrett's esophagus: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol.*, 16, pp1046-1055.
- Kuhar, N., Sil, S., Verma, T., Umaphathy, S. 2018, Challenges in application of Raman spectroscopy to biology and materials. *RSC Adv.*, 8, pp25888-25908.
- Kuipers, E.J., Spaander, M.C. 2018, Natural History of Barrett's Esophagus. *Dig Dis Sci.*, 63(8), pp1997-2004.

- Kunzmann, A.T., McMenamin, Ú.C., Spence, A.D., Gray, R.T, Murray, L.J., Turkington, R.C., Coleman, H.G. 2018, Blood Biomarkers for Early Detection of Oesophageal Cancer: A Systematic Review: A systematic review. *Eur. J. Gastroenterol. Hepatol.*, 30(3), pp263–273.
- Lagergren, J., Bergstrom, R., Lindgren, A., Nyrén, O. 1999, Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. *N Engl J Med.*, 340(11), pp825-31.
- Lai, H.S., Lee, J.C., Lee, P.H., Wang, S.T., Chen, W.J. 2005; Plasma free amino acid profile in cancer patients. *Semin Cancer Biol.*, 15, pp267–76.
- Layke, J.C., Lopez, P.P. 2006, Esophageal cancer: A Review and Update. *Am Fam Physician.*, 73(12), pp2187-2194.
- Leal, L.B., Nogueirab, M.S., Canevaria, R.A., Carvalho, L.F.C.S. 2018, Vibration spectroscopy and body biofluids: Literature review for clinical applications. *Photodiagnosis Photodyn Ther.*, 24, pp237-244.
- Leedham, S.J., Preston, S.L., Mcdonald, S.A., Elia, G., Bhandari, P., Poller, D., Harrison, R., Novelli, M.R., Jankowski, J.A., Wright, N.A. 2008, Individual crypt genetic heterogeneity and the origin of metaplastic glandular epithelium in human Barrett's oesophagus. *Gut.*, 57(8), pp 1041–1048.
- Levine, D.M., Ek, W.E., Zhang, R., Liu, X., Onstad, L., Sather, C., Lao-Sirieix, P., Gammon, M.D., Corley, D.A., Shaheen, N.J., Bird, N.C., Hardie, L.J., Murray, L.J., Reid, B.J., Chow, W.H., Risch, H.A., Nyrén, O., Ye, W., Liu, G., Romero, Y., Bernstein, L., Wu, A.H., Casson, A.G., Chanock, S.J., Harrington, P., Caldas, I., Debiram-Beecham, I., Caldas, C., Hayward, N.K., Pharoah, P.D., Fitzgerald, R.C., Macgregor, S., Whitman, D.C., Vaughan, T.L. 2013, A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus. *Nat Genet.*, 45(12), pp1487-1493.
- Li, Q., Sun, X., Xu, Y., Yang, L., Zhang, Y., Weng, S., Shi, J., Wu, J. 2005, Use Fourier-transform infrared Spectrosc. to rapidly diagnose gastric endoscopic. Biopsies. *World J Gastroenterol.*, 11, pp3842–3845.
- Li, X., Yang, T., Li, S. 2012, Discrimination of serum Raman spectroscopy between normal and colorectal cancer using selected parameters and regression-discriminant analysis. *Appl. Opt.*, 51, 5038.
- Li, Q, Hao, C., Kang, X., Zhang, J., Sun, X., Wang, W., Zeng, H. 2017, Colorectal Cancer and Colitis Diagnosis Using Fourier Transform Infrared Spectroscopy and an Improved K-Nearest-Neighbour Classifier. *Sensors.*, 17, pp1–9.
- Lomo, L.C., Blount, P.L., Sanchez, C.A., Li, X., Galipeau, P.C., Cowan, D.S., Ayub, K., Rabinovitch, P.S., Reid, B.J., Odze, R.D. 2006, Crypt dysplasia with surface maturation: a clinical, pathologic, and molecular study of a Barrett's esophagus cohort. *Am J Surg Pathol.*, 30(4), pp423-35.
- Longcroft-Wheaton, G., Duku, M., Mead, R., Poller, D., Bhandari, P. 2010, Acetic acid spray is an effective tool for the endoscopic detection of neoplasia in patients with Barrett's esophagus. *Clin Gastroenterol Hepatol.*, 8(10), pp843–7.
- Lovergne, L., Bounzy, P., Untereiner, V., Garnotel, R., Baker, M.J., Thiéfin, G., Sockalingum, G.D. 2016, Biofluid infrared spectro-diagnostics: pre-analytical considerations for clinical applications. *Faraday Discuss.*, 187, pp521–537.

- Lyng, F., Gaz, E., Gardner, P. 2011, Preparation of Tissues and Cells for Infrared and Raman Spectroscopy and Imaging in D. Moss (ed)Biomedical Applications of Synchrotron Infrared Microspectroscopy, *RSC Analytical Spectroscopy Monographs, No.*, 11, pp147-185.
- Macdonald, C.E., Wicks, A.C., Playford, R.J. 2000, Final results from 10-year cohort of patients undergoing surveillance for Barrett's oesophagus: observational study. *BMJ.*, 321, pp1252–1255.
- Maley, C.C., Galipeau, P.C., Finley, J.C., Wongsurawat, V.J., Li, X., Sanchez, C.A., Paulson, T.G., Blount, P.L., Risques, R.A., Rabinovitch, P.S., Reid, B.J. 2006, Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet.*, 38(4), pp468-473.
- Managò, S., Mirabelli, P., Napolitano, M., Zito, G., De Luca, A.C. 2018, Raman detection and identification of normal and leukemic hematopoietic cells. *J Biophotonics.*, 11(5), e201700265.
- Mannath J, Subramanian V, Hawkey CJ, Ragunath K. 2010, Narrow band imaging for characterization of high-grade dysplasia and specialized intestinal metaplasia in Barrett's esophagus: a meta-analysis. *Endoscopy.*, 42(5), pp351-9.
- Maziak, D.E., Do, M.T., Shamji, F.M., Sundaresan, S.R., Perkins, D.G., Wong, P.T. 2007, Fourier-transform infrared spectroscopic study of characteristic molecular structure in cancer cells of esophagus: an exploratory study., *Cancer Detect. Prev.*, 31, pp244–253.
- Meuse, C.W., Barker, P.E. 2009, Quantitative Infrared Spectroscopy of Formalin-fixed, Paraffin-embedded Tissue Specimens, Paraffin Wax Removal With Organic Solvents. *Appl. Immunohistochem. Mol. Morphol.*, 17(6), pp547–52.
- Meyskens, F.L., Mukhtar, H., Rock, C.L., Cuzick, J., Kensler, T.W, Yang, C.S., Ramsey, S.D., Lippman, S.M., Alberts, D.S. 2016, Cancer Prevention: Obstacles, Challenges, and the Road Ahead. *J Natl Cancer Inst.*, 108(2), djv309
- Mian, S., Colley, H., Thornhill, M., Martin, H., Rehman, I. 2014, Development of a Dewaxing Protocol for Tissue-Engineered Models of the Oral Mucosa Used for Raman Spectroscopic Analysis. *Applied Spectroscopy Reviews.*, 49, pp10.
- Mitchell, A.L., Gajjar, K.B., Theophilou, G., Martin, F.L., Martin-Hirsch, P.L. 2014, Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting. *J Biophotonics.*, 7(3-4), pp153-65.
- Montgomery, E., Bronner, M.P., Goldblum, J.R., Greenson, J.K., Haber, M.M., Hart, J., Lamps, L.W., Lauwers, G.Y., Lazenby, A.J., Lewin, D.N., Robert, M.E., Toledano, A.Y., Shyr, Y., Washington, K. 2001, Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol.*, 32, pp368–78.
- Morais, C.L.M., Lima, K.M.G. 2017, Comparing unfolded and two-dimensional discriminant analysis and support vector machines for classification of EEM data. *Chemometrics and Intelligent Laboratory Systems.*, 170, pp1-12.
- Morais, C.L.M., Lima, K.M.G. 2018, Principal Component Analysis with Linear and Quadratic Discriminant Analysis for Identification of Cancer Samples Based on Mass Spectrometry. *J. Braz. Chem. Soc.*, 29, pp472.

- Morais, C.L.M., Paraskevaidi, M., Cui, L., Fullwood, N.J., Isabelle, M., Lima, K.M.G., Martin-Hirsch, P.L., Sreedhar, H., Trevisan, J., Walsh, M.J., Zhang, D., Zhu, Y.G., Martin, F.L. 2019, Standardization of complex biologically derived spectrochemical datasets. *Nat. Protoc.*, 14, pp1546–1577.
- Morson, B.C., Belcher, J.R. 1952, Adenocarcinoma of the esophagus and ectopic gastric mucosa. *Br J Cancer.*, 6, pp127–130.
- Movasaghi, Z., Rehman, S., Rehman, I.U. 2007, Raman spectroscopy of Biological Tissues. *Appl. Spectrosc. Rev.*, 42, pp493–541.
- Moyes, L.H., Going, J.J. 2011, Still waiting for predictive biomarkers in Barrett's oesophagus. *J Clin Pathol.*, 64(9), pp742–750.
- Mudyadzo T.A. 2018, Barrett's Esophagus: A Molecular Overview. *Cureus.*, 10(10), e3468.
- National Institute for Health and Clinical Excellence (NICE). Gastro-oesophageal reflux disease and dyspepsia in adults: investigation and management. NICE, 2014. <https://www.nice.org.uk/guidance/cg184/chapter/1-Recommendations#surveillance-for-people-with-barretts-oesophagus-2> (accessed 7th November 2019).
- Nones, K., Waddell, N., Wayte, N., Patch, A.M., Bailey, P., Newell, F., Holmes, O., Fink, J.L., Quinn, M.C., Tang, Y.H., Lampe, G., Quek, K., Loffler, K.A., Manning, S., Idrisoglu, S., Miller, D., Xu, Q., Waddell, N., Wilson, P.J., Bruxner, T.J., Christ, A.N., Harliwong, I., Nourse, C., Nourbakhsh, E., Anderson, M., Kazakoff, S., Leonard, C., Wood, S., Simpson, P.T., Reid, L.E., Krause, L., Hussey, D.J., Watson, D.I., Lord, R.V., Nancarrow, D., Phillips, W.A., Gotley, D., Smithers, B.M., Whiteman, D.C., Hayward, N.K., Campbell, P.J., Pearson, J.V., Grimmond, S.M., Barbour, A.P. 2014, Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. *Nat Commun.*, 5, pp5224.
- Offman, J., Fitzgerald, R.C. 2017, Alternatives to traditional per-oral endoscopy for screening. *Gastrointest Endosc Clin N Am.*, 27, pp379–396.
- Old, O.J., Lloyd, G.R., Nallala, J., Isabelle, M., Almond, L.M., Shepherd, N.A., Kendall, C.A., Shore, A.C., Barr, H., Stone, N. 2017, Rapid infrared mapping for highly accurate automated histology in Barrett's oesophagus. *Analyst.*, 142(8), pp1227-1234.
- O'Riordan, J.M., Abdel-latif, M.M., Ravi, N., McNamara, D., Byrne, P.J., McDonald, G.S., Keeling, P.W., Kelleher, D., Reynolds, J.V. 2005, Proinflammatory cytokine and nuclear factor kappa-B expression along the inflammation–metaplasia–dysplasia–adenocarcinoma sequence in the esophagus. *Am J Gastroenterol.*, 100(6), pp1257-1264.
- Overholt, B.F., Wang, K.K., Burdick, J.S., Lightdale, C.J., Kimmey, M., Nava, H.R., Sivak, M.V. Jr, Nishioka, N., Barr, H., Marcon, N., Pedrosa, M., Bronner, M.P., Grace, M., Depot, M.; International Photodynamic Group for High-Grade Dysplasia in Barrett's Esophagus. 2007, Five-year efficacy and safety of photodynamic therapy with Photofrin in Barrett's high-grade dysplasia. *Gastrointest Endosc.*, 66(3), pp460-468.
- Paraskevaidi, M., Morais, C.L.M., Lima, K.M.G., Ashton, K.M., Stringfellow, H.F., Martin-Hirsch, P.L., Martin, F.L. 2018, Potential of mid-infrared spectroscopy as a non-invasive diagnostic test in urine for endometrial or ovarian cancer. *Analyst.*, 143(13), pp3156-3163.

- Paraskevaïdi, M., Martin-Hirsch, P.L., Martin, F.L. 2019, Need for early, minimally invasive cancer diagnosis. *Proc Natl Acad Sci U S A.*, 116(11), pp4752.
- Parker, F.S. 1983, Applications of infrared, Raman and Resonance Raman spectroscopy in biochemistry, vol. 9, Springer, New York.
- Petibois, C., Drogat, B., Bikfalvi, A., Dél  ris, G., Moenner M. 2007, Histological mapping of biochemical changes in solid tumors by FT-IR spectral imaging. *FEBS Lett.*, 581, pp5469–5474.
- Phoa, K.N., van Vilsteren, F.G., Weusten, B.L., Bisschops, R., Schoon, E.J., Rangunath, K., Fullarton, G., Di Pietro, M., Ravi, N., Visser, M., Offerhaus, G.J., Seldenrijk, C.A., Meijer, S.L., ten Kate, F.J., Tijssen, J.G., Bergman, J.J. 2014, Radiofrequency ablation vs endoscopic surveillance for patients with Barrett esophagus and low-grade dysplasia: a randomized clinical trial. *JAMA.*, 311(12), pp1209-1217.
- Pichardo-Molina, J.L., Frausto-Reyes, C., Barbosa-Garc  a, O., Huerta-Franco, R., Gonz  lez-Trujillo, J.L., Ram  rez-Alvarado, C.A., Guti  rrez-Ju  rez, G., Medina-Guti  rrez, C. 2007, Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients. *Lasers Med. Sci.*, 22, pp229–36.
- Playford, R.J. 2006, New British Society of Gastroenterology (BSG) guidelines for the diagnosis and management of Barrett’s oesophagus. *Gut.*, 55(4), pp442–3.
- Pohl, H., Welch, H.G. 2005, The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J Natl Cancer Inst.*, 97(2), pp142–146.
- Pohl, H., Pech, O., Arash, H., Stolte, M., Manner, H., May, A., Kraywinkel, K., Sonnenberg, A., Ell, C. 2016, Length of Barrett’s oesophagus and cancer risk: implications from a large sample of patients with early oesophageal adenocarcinoma. *Gut.*, 65(2), pp196-201.
- Pohl, J., Pech, O., May, A., Manner, H., Fissler-Eckhoff, A. 2010, Incidence of macroscopically occult neoplasias in Barrett’s esophagus: are random biopsies dispensable in the era of advanced endoscopic imaging? *Am J Gastroenterol.*, 105, pp2350–6.
- Pollit, V., Graham, D., Leonard, C., Filby, A., McMaster, J., Mealing, S.J., Lovat, L.B., Haidry, R.J. 2019, A cost-effectiveness analysis of endoscopic eradication therapy (EET) for management of dysplasia arising in patients with Barrett’s esophagus in the United Kingdom. *Curr Med Res Opin.*, 35(5), pp805-815.
- Prasad, G.A., Wang, K.K., Buttar, N.S., Wongkeesong, L.M., Krishnadath, K.K., Nichols, F.C. 3rd, Lutzke, L.S., Borkenhagen, L.S. 2007, Long-term survival following endoscopic and surgical treatment of high-grade dysplasia in Barrett’s esophagus. *Gastroenterology.*, 132(4), pp1226-1233.
- Provenzale, D., Schmitt, C., Wong, J.B. 1999, Barrett’s esophagus: a new look at surveillance based on emerging estimates of cancer risk. *Am J Gastroenterol.*, 94(8), pp2043–2053.
- Purandare, N.C., Trevisan, J., Patel, I.I., Gajjar, K., Mitchell, A.L., Theophilou, G., Valasoulis, G., Martin, M., von B  nau, G., Kyrgiou, M., Paraskevaïdis, E., Martin-Hirsch, P.L., Prendiville, W.J., Martin, F.L. 2013, Exploiting biospectroscopy as a novel screening tool for cervical cancer: towards a framework to validate its accuracy in a routine clinical setting. *Bioanalysis.*, 5, pp2697–711.

- Qiao, Y., Hyder, A., Bae, S.J., Zarin, W., O'Neill, T.J., Marcon, N.E., Stein, L., Thein, H.H. 2015, Surveillance in patients with Barrett's esophagus for early detection of esophageal adenocarcinoma: a systematic review and meta-analysis. *Clin Transl Gastroenterol.*, 10(6), e131.
- Quaroni, L., Casson, A.G. 2009, Characterization of Barrett esophagus and esophageal adenocarcinoma by Fourier-transform infrared microscopy. *Analyst.*, 134, pp1240–1246
- Qumseya, B.J., Wani, S., Gendy, S., Harnke, B., Bergman, J.J., Wolfsen, H. 2017, Disease progression in Barrett's low-grade dysplasia with radiofrequency ablation compared with surveillance: systematic review and meta-analysis. *Am J Gastroenterol.*, 112, pp849–865.
- Raman, C.V., Krishnan, K.S. 1928, A new type of secondary radiation. *Nature.*, 121, pp501–502.
- Rapado-González, O., Majem, B., Muínelo-Romay, L., López-López, R., Suarez-Cunqueiro, M.M. 2016, Cancer Salivary Biomarkers for Tumours Distant to the Oral Cavity. *Int. J. Mol Sci.*, 17, pp1531
- Reid, B.J., Prevo, L.J., Galipeau, P.C., Sanchez, C.A., Longton, G., Levine, D.S., Blount, P.L., Rabinovitch, P.S. 2001, Predictors of progression in Barrett's esophagus II: baseline 17p (*p53*) loss of heterozygosity identifies a patient subset at increased risk for neoplastic progression. *Am J Gastroenterol.*, 96(10), pp2839–2848.
- Robles, L.Y., Singh, S., Fisichella, P.M. 2015, Emerging enhanced imaging technologies of the esophagus: spectroscopy, confocal laser endomicroscopy and optical coherence tomography. *J Surg Research.*, 195, pp502-514.
- Rohleder, D.R., Kocherscheidt, G., Gerber, K., Kiefer, W., Köhler, W., Möcks, J., Petrich, W.H. 2005, Comparison of mid-infrared and Raman spectroscopy in the quantitative analysis of serum. *J. Biomedical Optics.*, 10(3), 031108.
- Ronkainen, J., Aro, P., Storskrubb, T., Johansson, S.E., Lind, T., Bolling-Sternevald, E., Vieth, M., Stolte, M., Talley, N.J., Agreus, L. 2005, Prevalence of Barrett's esophagus in the general population: An endoscopic study. *Gastroenterology.*, 129, pp1825–1831.
- Ross-Innes, C.S., Debiram-Beecham, I., O'Donovan, M., Walker, E., Varghese, S., Lao-Sirieix, P., Lovat, L., Griffin, M., Ragunath, K., Haidry, R., Sami, S.S., Kaye, P., Novelli, M., Disep, B., Ostler, R., Aigret, B., North, B.V., Bhandari, P., Haycock, A., Morris, D., Attwood, S., Dhar, A., Rees, C., Rutter, M.D., Sasieni, P.D., Fitzgerald, R.C.; BEST2 Study Group. 2015, Evaluation of a minimally invasive cell sampling device coupled with assessment of trefoil factor 3 expression for diagnosing Barrett's esophagus: a multi-center case–control study. *PLoS Med.*, 12(1), e1001780.
- Ruol, A., Parenti, A., Zaninotto, G. 2000, Intestinal metaplasia is the probable common precursor of adenocarcinoma in barrett esophagus and adenocarcinoma of the gastric cardia. *Cancer.*, vol. 88(11), pp2520–2528.
- Saad, R.S., Mahood, L.K., Clary, K.M. 2003, Role of cytology in the diagnosis of Barrett's esophagus and associated neoplasia. *Diagn Cytopathol.*, 29, pp130–135.
- Sahu, A., Sawant, S., Mamgain, H., Krishna, C.M. 2013, Raman spectroscopy of serum: an exploratory study for detection of oral cancers. *Analyst.*, 138, 4161–4174.

- Sahu, R.K., Argov, S., Walfisch, S., Bogomolny, E., Moreh, R., Mordechai, S. 2010, Prediction potential of IR-micro spectroscopy for colon cancer relapse. *Analyst.*, 135(3), pp538-44.
- Sahu, R.K., Salman, A., Mordechai, S. 2017. Tracing overlapping biological signals in mid-infrared using colonic tissues as a model system. *World J Gastroenterol.*, 23(2), pp286-296.
- Sala, A., Anderson, D.J., Brennan, P.M., Holly J Butler, H.J., Cameron, J.M., Jenkinson, M.D., Rinaldi, C., Theakstone, A.G., Baker, M.J. 2020, Biofluid diagnostics by FTIR spectroscopy: A platform technology for cancer detection. *Cancer Lett.*, 477, pp122-130.
- Sampliner, R.E. 1998, Practice guidelines on the diagnosis, surveillance, and therapy of Barrett's esophagus: the Practice Parameters Committee of the American College of Gastroenterology. *Am J Gastroenterol.*, 93, pp1028–1032.
- Sanghi, V., Thota, P.N. 2019, Barrett's esophagus: novel strategies for screening and surveillance. *Ther Adv Chronic Dis.*, 10, 2040622319837851
- Sarr, M.G., Hamilton, S.R., Marrone, G.C., Cameron, J.L. 1985, Barrett's esophagus: It's prevalence and association with adenocarcinoma in patients with symptoms of gastroesophageal reflux. *Am J Surg.*, 149, pp187–93
- Saxena, N., Inadomi, J.M. 2017, Effectiveness and cost-effectiveness of endoscopic screening and surveillance. *Gastrointest Endosc Clin N Am.*, 27, pp397–421.
- Schlemper, R.J., Riddell, R.H., Kato, Y., Borchard, F., Cooper, H.S., Dawsey, S.M., Dixon, M.F., Fenoglio-Preiser, C.M., Fléjou, J.F., Geboes, K., Hattori, T., Hirota, T., Itabashi, M., Iwafuchi, M., Iwashita, A., Kim, Y.I., Kirchner, T., Klimpfinger, M., Koike, M., Lauwers, G.Y., Lewin, K.J., Oberhuber, G., Offner, F., Price, A.B., Rubio, C.A., Shimizu, M., Shimoda, T., Sipponen, P., Solcia, E., Stolte, M., Watanabe, H., Yamabe, H. 2000, The Vienna classification of gastrointestinal epithelial neoplasia. *Gut.*, 47(2), pp251–5.
- Shaheen, N.J., Crosby, M.A., Bozyski, E.M., Sandler, R.S. 2000, Is there publication bias in the reporting of cancer risk in Barrett's esophagus? *Gastroenterology.*, 119, pp333–8.
- Shapiro, B., Chakrabarty, M., Cohn, E.M., Leon, S.A. 1983, Determination of circulating DNA levels in patients with benign or malignant gastrointestinal disease. *Cancer.*, 51(11), pp2116-20.
- Shah, P.M., Gerdes, H. 2015, Endoscopic options for early stage esophageal cancer. *J Gastrointest Oncol.*, 6(1), pp20–30.
- Sharma, P., Dent, J., Armstrong, D., Bergman, J.J., Gossner, L., Hoshihara, Y., Jankowski, J.A., Junghard, O., Lundell, L., Tytgat, G.N., Vieth, M. 2006, The development and validation of an endoscopic grading system for Barrett's esophagus: the Prague C & M criteria. *Gastroenterology.*, 131, pp1392-9. (A)
- Sharma, P., Marcon, N., Wani, S., Bansal, A., Mathur, S., Sampliner, R., Lightdale, C. 2006, Non-biopsy detection of intestinal metaplasia and dysplasia in Barrett's esophagus: a prospective multicenter study. *Endoscopy.*, 38(12), pp1206–12. (B)

- Sharma, P., Morales, T.G., Sampliner, R.E. 1998, Short segment Barrett's esophagus—the need for standardization of the definition and of endoscopic criteria. *Am J Gastroenterology.*, 93(7), pp1033–1036.
- Sharma, P., Weston, A.P., Morales, T., Topalovski, M., Mayo, M.S., Sampliner, R.E. 2000, Relative risk of dysplasia for patients with intestinal metaplasia in the distal oesophagus and in the gastric cardia. *Gut.*, 46(1), pp9–13.
- Sharma, P., Weston, A.P., Topalovski, M., Cherian, R., Bhattacharyya, A., Sampliner, R.E. 2003, Magnification chromoendoscopy for the detection of intestinal metaplasia and dysplasia in Barrett's oesophagus. *Gut.*, 52, pp24–7.
- Shariff, M.K., Bird-Lieberman, E.L., O'Donovan, M., Abdullahi, Z., Liu, X., Blazeby, J., Fitzgerald, R.C. 2012, Randomized crossover study comparing efficacy of transnasal endoscopy with that of standard endoscopy to detect Barrett's esophagus. *Gastrointest Endosc.*, 75(5), pp954-961.
- Shiota, S., Singh, S., Anshasi, A., El-Serag, H.B. 2015, Prevalence of Barrett's esophagus in Asian countries: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol.*, 13, pp1907-1918.
- Singh, S., Manickam, P., Amin, A.V., Samala, N., Schouten, L.J., Iyer, P.G., Desai, T.K. 2014, Incidence of esophageal adenocarcinoma in Barrett's esophagus with low-grade dysplasia: a systematic review and meta-analysis. *Gastrointest Endosc.*, 79, pp897-909.
- Skacel, M., Petras, R.E., Rybicki, L.A., Gramlich, T.L., Richter, J.E., Falk, G.W., Goldblum, J.R. 2002, *p53* expression in low grade dysplasia in Barrett's esophagus: correlation with interobserver agreement and disease progression. *Am J Gastroenterol.*, 97, pp2508–13.
- Slaughter, D.P., Southwick, H.W., Smejkal, W. 1953, Field cancerization" in oral stratified squamous epithelium. *Cancer (Phila.)*, 6, pp963-968.
- Small, A.J., Araujo, J.L., Leggett, C.L., Mendelson, A.H., Agarwalla, A., Abrams, J.A., Lightdale, C.J., Wang, T.C., Iyer, P.G., Wang, K.K., Rustgi, A.K., Ginsberg, G.G., Forde, K.A., Gimotty, P.A., Lewis, J.D., Falk, G.W., Bewtra, M. 2015, Radiofrequency ablation is associated with decreased neoplastic progression in patients with Barrett's esophagus and confirmed low-grade dysplasia. *Gastroenterology.*, 149(3), pp567-576.e3
- Somja, J., Demoulin, S., Roncarati, P., Herfs, M., Bletard, N., Delvenne, P., Hubert, P. 2013, Dendritic cells in Barrett's esophagus carcinogenesis: an inadequate microenvironment for antitumor immunity? *Am J Pathol.*, 182(6), pp2168–2179.
- Spechler, S.J., Sharma, P., Souza, R.F., Inadomi, J.M., Shaheen, N.J. 2011, American gastroenterological association technical review on the management of Barrett's esophagus. *Gastroenterology.*, 140(3), ppe18–e52.
- Srivastava, A., Hornick, J.L., Li, X., Blount, P.L., Sanchez, C.A., Cowan, D.S., Ayub, K., Maley, C.C., Reid, B.J., Odze, R.D. 2007, Extent of low-grade dysplasia is a risk factor for the development of esophageal adenocarcinoma in Barrett's esophagus. *Am J Gastroenterol.*, 102(3), pp483-493.
- Stachler, M.D., Taylor-Weiner, A., Peng, S., Mckenna, A., Agoston, A.T., Odze, R.D., Davison, J.M., Nason, K.S., Loda, M., Leshchiner, I., Stewart, C., Stojanov, P., Seepo, S., Lawrence, M.S., Ferrer-Torres, D., Lin, J., Chang, A.C., Gabriel, S.B., Lander, E.S., Beer, D.G., Getz, G., Carter, S.L., Bass,

A.J. 2015, Paired exome analysis of Barrett's esophagus and adenocarcinoma. *Nat Genet.*, 47(9), pp1047–1055.

Stewart, S., Priore, R.J., Nelson, M.P., Treado, P.J. 2012, Raman imaging. *Annu. Rev. Anal. Chem.*, 5, pp337–360.

Stone, N., Kendall, C., Shepherd, N., Crow, P., Barr, H. 2004, Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *Faraday Discuss.*, 126, pp141-57; discussion pp169-83.

Su, Z., Gay, L.J., Strange, A., Palles, C., Band, G., Whiteman, D.C., Lescai, F., Langford, C., Nanji, M., Edkins, S., van der Winkel, A., Levine, D., Sasieni, P., Bellenguez, C., Howarth, K., Freeman, C., Trudgill, N., Tucker, A.T., Pirinen, M., Peppelenbosch, M.P., van der Laan, L.J., Kuipers, E.J., Drenth, J.P., Peters, W.H., Reynolds, J.V., Kelleher, D.P., McManus, R., Grabsch, H., Prenen, H., Bisschops, R., Krishnadath, K., Siersema, P.D., van Baal, J.W., Middleton, M., Petty, R., Gillies, R., Burch, N., Bhandari, P., Paterson, S., Edwards, C., Penman, I., Vaidya, K., Ang, Y., Murray, I., Patel, P., Ye, W., Mullins, P., Wu, A.H., Bird, N.C., Dallal, H., Shaheen, N.J., Murray, L.J., Koss, K., Bernstein, L., Romero, Y., Hardie, L.J., Zhang, R., Winter, H., Corley, D.A., Panter, S., Risch, H.A., Reid, B.J., Sargeant, I., Gammon, M.D., Smart, H., Dhar, A., McMurtry, H., Ali, H., Liu, G., Casson, A.G., Chow, W.H., Rutter, M., Tawil, A., Morris, D., Nwokolo, C., Isaacs, P., Rodgers, C., Ragnath, K., MacDonald, C., Haigh, C., Monk, D., Davies, G., Wajed, S., Johnston, D., Gibbons, M., Cullen, S., Church, N., Langley, R., Griffin, M., Alderson, D., Deloukas, P., Hunt, S. E., Gray, E., Dronov, S., Potter, S.C., Tashakkori-Ghanbaria, A., Anderson, M., Brooks, C., Blackwell, J.M., Bramon, E., Brown, M.A., Casas, J.P., Corvin, A., Duncanson, A., Markus, H.S., Mathew, C.G., Palmer, C.N., Plomin, R., Rautanen, A., Sawcer, S.J., Trembath, R.C., Viswanathan, A.C., Wood, N., Trynka, G., Wijmenga, C., Cazier, J.B., Atherfold, P., Nicholson, A.M., Gellatly, N.L., Glancy, D., Cooper, S.C., Cunningham, D., Lind, T., Hapeshi, J., Ferry, D., Rathbone, B., Brown, J., Love, S., Attwood, S., MacGregor, S., Watson, P., Sanders, S., Ek, W., Harrison, R.F., Moayyedi, P., de Caestecker, J., Barr, H., Stupka, E., Vaughan, T.L., Peltonen, L., Spencer, C.C., Tomlinson, I., Donnelly, P., Jankowski, J.A.; Esophageal Adenocarcinoma Genetics Consortium; Wellcome Trust Case Control Consortium 2. 2012, Common variants at the MHC locus and at chromosome 16q24.1 predispose to Barrett's esophagus. *Nat Genet.*, 44(10), pp1131-1136.

Takubo, K., Vieth, M., Aida, J., Sawabe, M., Kumagai, Y., Hoshihara, Y., Arai, T. 2009, Differences in the definitions used for esophageal and gastric diseases in different countries: endoscopic definition of the esophagogastric junction, the precursor of Barrett's adenocarcinoma, the definition of Barrett's esophagus, and histologic criteria for mucosal adenocarcinoma or high-grade dysplasia. *Digestion.*, 80(4), pp248–257.

Talari, A.C.S., Rehman, S., Rehman, I.U. 2019, Advancing cancer diagnostics with artificial intelligence and spectroscopy: identifying chemical changes associated with breast cancer. *Expert. Rev. Mol. Diagn.*, 19, pp929.

Taleb, I., Thiéfin, G., Gobinet, C., Untereiner, V., Bernard-Chabert, B., Heurgué, A., Truntzer, C., Hillon, P., Manfait, M., Ducoroy, P., Sockalingum, G.D. 2013, Diagnosis of hepatocellular carcinoma in cirrhotic patients: a proof-of-concept study using serum micro-Raman spectroscopy. *Analyt.*, 138(14), pp4006-14.

- Theophilou, G., Lima, K.M., Briggs, M., Martin-Hirsch, P.L., Stringfellow, H.F., Martin, F.L. 2015, A biospectroscopic analysis of human prostate tissue obtained from different time periods points to a trans-generational alteration in spectral phenotype. *Sci. Rep.*, 27(5), pp13465.
- Theophilou, G., Lima, K.M.G., Martin-Hirsch, P.L., Stringfellow, H.F., Martin F.L. 2016, ATR-FTIR spectroscopy coupled with chemometric analysis discriminates normal, borderline and malignant ovarian tissue: classifying subtypes of human cancer. *Analyst.*, 141(2), pp585-94.
- Thomas, P., Doddoli, C., Lienne, P., Morati, N., Thirion, X., Garbe, L., Giudicelli, R., Fuentes, P. 1997, Changing patterns and surgical results in adenocarcinoma of the oesophagus. *Br J Surg.*, 84, pp119–25.
- Thomas, T., Abrams, K.R., De Caestecker, J.S., Robinson, R.J. 2007, Meta-analysis: cancer risk in Barrett's oesophagus. *Aliment Pharmacol Ther.*, 26, pp1465–1477.
- Timmer, M.R., Sun, G., Gorospe, E.C., Leggett, C.L., Lutzke, L., Krishnath, K.K., Wang, K.K. 2013, Predictive biomarkers for Barrett's esophagus: so near and yet so far. *Dis Esophagus.*, 26(6), pp574–581.
- Trivedi, P.J., Brade, B. 2013, Indications, stains and techniques in chromoendoscopy. *QJM.*, 106(2), pp117–131.
- Underwood, T.J., Hayden, A.L., Derouet, M., Garcia, E., Noble, F., White, M.J., Thirdborough, S., Mead, A., Clemons, N., Mellone, M., Uzoho, C., Primrose, J.N., Blaydes, J.P., Thomas, G.J. 2015, Cancer-associated fibroblasts predict poor outcome and promote periostin-dependent invasion in oesophageal adenocarcinoma. *J Pathol.*, 235(3), pp466–477.
- van Blankenstein, M., Looman, C.W., Johnston, B.J., Caygill, C.P. 2005, Age and sex distribution of the prevalence of Barrett's esophagus found in a primary referral endoscopy center. *Am J Gastroenterol.*, 100, pp568–576.
- Vazquez-Zapien, G.J., Mata-Miranda, M.M., Sanchez-Monroy, V., Delgado-Macuil, R.J., Perez-Ishiwara, D.G., Rojas-Lopez, M. 2016, FTIR Spectroscopic and Molecular Analysis during Differentiation of Pluripotent Stem Cells to Pancreatic Cells. *Stem Cells Int.*, 2016, 2016, pp1–10.
- Virkler, K., Lednev, I.K. 2010, Forensic body fluid identification: the Raman spectroscopic signature of saliva. *Analyst.*, 135, pp512–517.
- Wajed, S.A., Laird, P.W., DeMeester, T.R. 2001, DNA methylation: an alternative pathway to cancer. *Ann. Surg.*, 234, pp10.
- Wang, J.S., Xu, Y.Z., Shi, J.S., Zhang, L., Duan, X.Y., Yang, L.M., Su, Y.L., Weng, S.F., Xu, D.F., Wu, J.G. 2003, FTIR spectroscopic study on normal and cancerous tissues of esophagus. *Guangpuxue Yu Guangpu Fenxi.*, 23, pp863–865.
- Wang, T.D., Triadafilopoulos, G., Crawford, J.M., Dixon, L.R., Bhandari, T., Sahbaie, P., Friedland, S., Soetikno, R., Contag, C.H. 2007, Detection of endogenous biomolecules in Barrett's esophagus by Fourier transform infrared spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.*, 104, pp15864–15869.
- Wang, K.K., Sampliner, R.E. 2008, Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *Am J Gastroenterol.*, 103(3), pp788–797.

- Weaver, J.M.J., Ross-Innes, C.S., Shannon, N., Lynch, A.G., Forshe, T., Barbera, M., Murtaza, M., Ong, C.J., Lao-Sirieix, P., Dunning, M.J., Smith, L., Smith, M.L., Anderson, C.L., Carvalho, B., O'Donovan, M., Underwood, T.J., May, A.P., Grehan, N., Hardwick, R., Davies, J., Oloumi, A., Aparicio, S., Caldas, C., Eldridge, M.D., Edwards, P.A.W., Rosenfeld, N., Tavaré, S., Fitzgerald, R.C.; OCCAMS consortium. 2014, Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nat Genet.*, 46(8), pp:837-843.
- Wiens, R., Rak, M., Cox, N., Abraham, S., Juurlink, B.H.J., Kulyk, W.M., Gough, K.M. 2007, Synchrotron FTIR microspectroscopic analysis of the effects of anti-inflammatory therapeutics on wound healing in laminectomized rats. *Anal. Bioanal. Chem.*, 387, pp1679–1689.
- Witt, T. R., Bains, M. S., Zaman, M. B., Martini, N. 1983. Adenocarcinoma in Barrett's esophagus. *J Thorac Cardiovasc Surg.*, 85(3), pp337-345.
- Wong, A., Fitzgerald, R.C. 2005, Epidemiologic risk factors for Barrett's esophagus and associated adenocarcinoma. *Clin Gastroenterol Hepatol.* 3, pp1–10.
- Wu, M., Zhang, W., Tian, P., Ling, X., Xu, Z. 2016, Intraoperative diagnosis of thyroid diseases by fourier transform infrared spectroscopy based on support vector machine. *Int. J Clin. Exp. Med.*, 9, pp2351–2358.
- Xu, Y.Z., Soloway, R.D., Lin, X.F., Zhi, X., Weng, S.F., Wu, Q.G., Shi, J.S., Sun, W.X., Zhang, T.X., Wu, J.G., Xu, D.F, Xu, G.X. 2000, Fourier transform infrared (FT-IR) mid-IR spectroscopy separates normal and malignant tissue from the colon and stomach. *Gastroenterology.*, 118, A6438
- Yachimski, P., Lee, R.A., Tramontano, A., Nishioka, N.S., Hur, C. 2010, Secular trends in patients diagnosed with Barrett's esophagus, *Dig Dis Sci.*, 55(4), pp960–966.
- Zagari, R.M., Fuccio, L., Wallander, M.A., Johansson, S., Fiocca, R., Casanova, S., Farahmand, B.Y., Winchester, C.C., Roda, E., Bazzoli, F. 2008, Gastro-oesophageal reflux symptoms, oesophagitis and Barrett's oesophagus in the general population: the Loiano-Monghidoro study. *Gut*, 57, pp1354–1359.

10. Appendix

10.1 Supplementary information (SI)

Table S1: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using plasma samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm⁻¹)	Tentative Assignments
929	Left-handed helix DNA (Z form)
952	Symmetric stretching vibration of $\nu 1\text{PO}_4^{3-}$ (phosphate of HA)
987	OCH ₃ (polysaccharides-cellulose)
999	Ring stretching vibrations mixed strongly with CH in-plane bending
1018	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1053	$\nu\text{C-O}$ & $\delta\text{C-O}$ of carbohydrates
1072	Phosphate I band for two different C-O vibrations of Deoxyribose in DNA in disordering structure
1134	Oligosaccharide C-OH stretching band 2-Methylmannoside
1180	Amide III band region
1354	Stretching C-O, deformation C-H, deformation N-H
1381	δCH_3 Stretching C-O, deformation C-H, deformation N-H
1388	Carbon particle
1392	Less characteristic, due to aliphatic side groups of the amino acid residues
1431	$\delta(\text{CH}_2)$ (polysaccharides, cellulose)
1485	Deformation C-H
1539	Protein amide II absorption- predominately β -sheet of amide II
1585	Ring C-C stretch of phenyl
1624	Amide I region
1643	Amide I band (arises from C=O stretching vibrations)
1681	C=O Guanine deformation N-H in plane
1712	C=O
1724	C=O stretching band mode of the fatty acid ester

Table S2: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using saliva samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
902	Phosphodiester region
991	C-O ribose, C-C
1003	Ring stretching vibrations mixed strongly with CH in-plane bending
1014	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1068	Stretching C-O ribose
1099	Stretching PO_2^- symmetric (phosphate II)
1107	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1431	$\delta(\text{CH}_2)$ (polysaccharides, cellulose)
1558	Ring base
1585	Ring C-C stretch of phenyl
1589	Ring C-C stretch of phenyl
1604	Adenine vibration in DNA
1624	Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode
1643	Amide I band (arises from C=O stretching vibrations)
1689	Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode
1697	C ₂ =O guanine
1701	C ₅ =O guanine
1716	C=O thymine
1743	C=O stretching mode of lipids
1778	$\nu(\text{C}=\text{C})$ lipids, fatty acids
1786	$\nu(\text{C}=\text{C})$ lipids, fatty acids

Table S3: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using serum samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
1041	Symmetric PO ₂ ⁻ stretching in RNA and DNA
1315	Amide III band components of proteins
1319	Amide III band components of proteins Collagen
1330	CH ₂ wagging
1338	CH ₂ wagging
1435	δ (CH ₂) (polysaccharides, cellulose)
1442	δ(CH ₂), lipids, fatty acids
1446	δ(CH ₂), lipids, fatty acids
1477	CH ₂ bending of the methylene chains in lipids
1492	C=C, deformation C-H
1539	Amide II
1573	C=N adenine
1593	Ring C-C stretch of phenyl
1600	Amide I band of proteins
1631	Ring C-C stretch of phenyl
1654	Amide I (of proteins in α-helix conformation)
1662	Amide I band
1743	C=O stretching mode of lipids

Table S4: Category-distinguishing wavenumbers for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using urine samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
906	Phosphodiester region
956	C-O deoxyribose, C-C
995	Ring breathing
1018	Glycogen
1030	Glycogen
1095	$\nu_{\text{as}}\text{PO}_2^-$
1118	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1141	Symmetric PO_2^- stretching in RNA and DNA
1242	Amide III collagen
1253	Amide III
1334	$\delta(\text{CH})$, ring (polysaccharides, pectin)
1381	δCH_3
1431	Lipids
1446	$\delta(\text{CH}_2)$, lipids, fatty acids
1500	In-plane CH bending vibration from the phenyl rings
1550	Amide II of proteins
1562	Amide II of proteins (<i>e.g.</i> , side-chain carboxyl groups)
1577	Glutamate carboxylate stretching
1600	C=N cytosine, N-H adenine
1651	Amide I
1681	C=O Guanine deformation N-H in plane
1712	C=O thymine
1724	C=O stretching band mode of the fatty acid ester
1735	C=O stretching (lipids)
1777	$\nu(\text{C}=\text{C})$ lipids, fatty acids

Table S5: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using plasma samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
929	Left-handed helix DNA (Z form)
952	Symmetric stretching vibration of $\nu_1\text{PO}_4^{3-}$ (phosphate of HA)
987	OCH ₃ (polysaccharides-cellulose)
999	Ring stretching vibrations mixed strongly with CH in-plane bending
1018	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1053	$\nu\text{C-O}$ & $\delta\text{C-O}$ of carbohydrates
1072	Phosphate I band for two different C-O vibrations of Deoxyribose in DNA in disordering structure
1134	Oligosaccharide C-OH stretching band 2-Methylmannoside
1180	Amide III band region
1354	Stretching C-O, deformation C-H, deformation N-H
1381	δCH_3 Stretching C-O, deformation C-H, deformation N-H
1388	Carbon particle
1392	Less characteristic, due to aliphatic side groups of the amino acid residues
1431	$\delta(\text{CH}_2)$ (polysaccharides, cellulose)
1485	Deformation C-H
1539	Protein amide II absorption- predominately β -sheet of amide II
1585	Ring C-C stretch of phenyl
1624	Amide I region
1643	Amide I band (arises from C=O stretching vibrations)
1681	C=O Guanine deformation N-H in plane
1712	C=O
1724	C=O stretching band mode of the fatty acid ester

Table S6: Category-distinguishing wavenumbers for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using saliva samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
902	Phosphodiester region
991	C-O ribose, C-C
1003	Ring stretching vibrations mixed strongly with CH in-plane bending
1014	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1068	Stretching C-O ribose
1099	Stretching PO_2^- symmetric (phosphate II)
1107	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1431	$\delta(\text{CH}_2)$ (polysaccharides, cellulose)
1558	Ring base
1585	Ring C-C stretch of phenyl
1589	Ring C-C stretch of phenyl
1604	Adenine vibration in DNA
1624	Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode
1643	Amide I band (arises from C=O stretching vibrations)
1689	Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode
1697	C ₂ =O guanine
1701	C5=O guanine
1716	C=O thymine
1743	C=O stretching mode of lipids
1778	$\nu(\text{C}=\text{C})$ lipids, fatty acids
1786	$\nu(\text{C}=\text{C})$ lipids, fatty acids

Table S7: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using serum samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
1041	Symmetric PO ₂ ⁻ stretching in RNA and DNA
1315	Amide III band components of proteins
1319	Amide III band components of proteins Collagen
1330	CH ₂ wagging
1338	CH ₂ wagging
1435	δ (CH ₂) (polysaccharides, cellulose)
1442	δ(CH ₂), lipids, fatty acids
1446	δ(CH ₂), lipids, fatty acids
1477	CH ₂ bending of the methylene chains in lipids
1492	C=C, deformation C-H
1539	Amide II
1573	C=N adenine
1593	Ring C-C stretch of phenyl
1600	Amide I band of proteins
1631	Ring C-C stretch of phenyl
1654	Amide I (of proteins in α-helix conformation)
1662	Amide I band
1743	C=O stretching mode of lipids

Table S8: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using urine samples obtained for SPA-QDA and GA-QDA models.

Wavenumbers (cm ⁻¹)	Tentative Assignments
906	Phosphodiester region
956	C-O deoxyribose, C-C
995	Ring breathing
1018	Glycogen
1030	Glycogen
1095	$\nu_{\text{as}}\text{PO}_2^-$
1118	$\nu(\text{CO})$, $\nu(\text{CC})$, $\delta(\text{OCH})$, ring (polysaccharides, pectin)
1141	Symmetric PO_2^- stretching in RNA and DNA
1242	Amide III collagen
1253	Amide III
1334	$\delta(\text{CH})$, ring (polysaccharides, pectin)
1381	δCH_3
1431	Lipids
1446	$\delta(\text{CH}_2)$, lipids, fatty acids
1500	In-plane CH bending vibration from the phenyl rings
1550	Amide II of proteins
1562	Amide II of proteins (<i>e.g.</i> , side-chain carboxyl groups)
1577	Glutamate carboxylate stretching
1600	C=N cytosine, N-H adenine
1651	Amide I
1681	C=O Guanine deformation N-H in plane
1712	C=O thymine
1724	C=O stretching band mode of the fatty acid ester
1735	C=O stretching (lipids)
1777	$\nu(\text{C}=\text{C})$ lipids, fatty acids

Table S9: Category-distinguishing wavenumbers for normal *vs.* inflammatory *vs.* Barrett's oesophagus *vs.* LGD *vs.* HGD *vs.* OAC using spectra tissue samples obtained for SPA-QDA model.

Wavenumbers (cm⁻¹)	Tentative Assignments
902	Phosphodiester region
952	Symmetric stretching vibration of phosphate
995	Ring breathing
1022	Glycogen
1072	Nucleic acid band
1111	Carbohydrates
1157	C-O stretching vibration
1226	phosphate I
1300	N-H thymine
1338	CH ₂ wagging
1361	Stretching C-O
1377	Stretching C-N cytosine
1392	Carbon particle
1431	polysaccharides
1450	Methylene deformation in biomolecules
1462	CH ₂ scissoring mode of the acyl chain of lipid
1485	Deformation C-H
1504	In-plane CH bending vibration from the phenyl rings
1539	Protein amide II absorption
1550	Amide II
1570	Amide II
1600	C=O stretching (lipids)
1624	Ring C-C stretch of phenyl
1643	Amide I band
1651	Amide I absorption
1662	Amide I band
1697	N-H thymine
1724	C=O stretching band mode of the fatty acid ester
1747	C=O stretching mode of lipids
1800	fatty acids

Table S10: Category-distinguishing wavenumbers for normal vs. inflammatory vs. Barrett's oesophagus vs. LGD vs. HGD vs. OAC using spectra tissue samples obtained for GA-QDA models.


Wavenumbers (cm⁻¹)	Tentative Assignments
884	Proteins, including collagen I
1188	Anti-symmetric phosphate vibrations
1206	Amide III (proteins)
1235	Amide III
1296	CH ₂ deformation
1307	CH ₃ /CH ₂ twisting or bending mode of lipid/collagen
1365	Tryptophan
1383	CH ₃ band
1402	Bending modes of methyl groups (one of vibrational modes of collagen)
1440	CH ₂ and CH ₃ deformation vibrations
1461	CH ₂ /CH ₃ deformation of lipids & collagen
1608	Tyrosine, phenylalanine ring vibration
1641	Amide I band (protein band)
1656	Amide I (proteins)
1715	C=O
1793	C=O


Royal Preston Hospital
 Sharoe Green Lane
 Fulwood
 Preston
 PR2 9HT


PATIENT CONSENT FORM


Title of Project: Collection of Gastro-intestinal Tissue and/or Blood, Saliva & Urine Samples for Future Research Use – Cancer Research Tissue Bank

Name of Researchers: Dr M Pitt, Mr Ravindra Date, Mr K Pursnani, Miss V Shetty, Mr J Ward, Mr P Turner, Mr C Ball


• I confirm that I have read and understand the information sheet (Version 4, 07.01.16) for the above tissue collection and have had the opportunity to ask questions. 


• I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected. 


• I understand that samples will be collected during diagnosis, surgery, at post-operative appointments and during any subsequent routine appointments. 

• I understand that sections of any of my medical notes from Lancashire Teaching Hospitals or elsewhere may be looked at and information taken from them to be analysed in strict confidence by responsible individuals from the research team or from regulatory authorities where it is relevant to my taking part in research. I give permission for these individuals to have access to my records. 

5. I agree to take part in the above study.

6. **Consent for storage and use in possible future research projects:** I agree that the samples I have given and the information gathered about me can be stored by the Pathology Laboratory at Lancashire Teaching Hospitals for possible use in future projects, subject to additional project specific ethical approval. 

7. **Genetic research:** I understand that future approved projects utilising the sample(s) I have given, may include genetic research aimed at understanding the genetic influences on tumours, but that the results of these investigations are unlikely to have any implications for me personally. 

8. **Future clinical findings:** I understand that in the short term there are unlikely to be findings that will be of direct relevance to me, however I would like to be informed in the future should this change 

Please sign and date

 Name of Patient

 Date

 Signature

 Name of Person taking declaration

 Date

 Signature



Health Research Authority

East of England - Cambridge Central Research Ethics Committee

Royal Standard Place
Nottingham
NG1 6FS

07 June 2018

Dr Tweesha Wahie
Lancashire Teaching Hospitals NHS Trust
Pathology Department
Lancashire Teaching Hospitals
RPH, Sharoe Green Lane, Fulwood, Preston
PR2 9HT

Dear Dr Wahie

Title of the Research Tissue Bank: Archival gastro-intestinal tissue, blood, saliva and urine collection
REC reference: 18/EE/0069
Designated Individual: Prof Timothy Dawson
IRAS project ID: 242639

Thank you for your letter of 17/04/2018, responding to the Committee's request for further information on the above research tissue bank and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the Research Tissue Bank on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all Research Tissue Banks that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact hra.studyregistration@nhs.net outlining the reasons for your request. Under very limited circumstances (e.g., for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the Research Tissue Bank.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion of the above research tissue bank on the basis described in the application form and supporting documentation as revised.

The Committee has also confirmed that the favourable ethical opinion applies to all research projects conducted in the UK using tissue or data supplied by the tissue bank, provided that the release of tissue or data complies with the attached conditions. It will not be necessary for these researchers to make project-based applications for ethical approval.

Duration of ethical opinion

The favourable opinion is given for a period of five years from the date of this letter and provided that you comply with the standard conditions of ethical approval for Research Tissue Banks set out in the attached document. You are advised to study the conditions carefully. The opinion may be renewed for a further period of up to five years on receipt of a fresh application. It is suggested that the fresh application is made 3-6 months before the 5 years expires, to ensure continuous approval for the research tissue bank.

Approved documents

The documents reviewed and approved at the meeting were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Human Tissue Authority licence [Scan of Licence]	N/A	27 August 2010
Other [Management of Tissue Bank Samples]	16.1	29 November 2016
Other [CI Tweesha Wahie]	N/A	30 January 2018
Other [Trust consent form printable]	Current	26 January 2018
Other [Response Letter]	N/A	17 April 2018
Other [Small Animal letter]	N/A	19 March 2015
Participant consent form [Patient consent form]	5	03 April 2018
Participant information sheet (PIS) [Patient information sheet]	5	03 April 2018
Protocol for management of the tissue bank [Gastro-intestinal tissue bank protocol]	5	25 January 2018
REC Application Form [RTB_Form_02022018]	4.1	02 February 2018

Research governance

A copy of this letter is being sent to the R&D office responsible for Lancashire Teaching Hospitals NHS Trust. You are advised to check their requirements for approval of the research tissue bank.

Under the Research Governance Framework (RGF), there is no requirement for NHS research permission for the establishment of research tissue banks in the NHS. Applications to NHS R&D offices through IRAS are not required as all NHS organisations are expected to have included management review in the process of establishing the research tissue bank.

Research permission is also not required by collaborators at tissue collection centres (TCCs) who provide tissue or data under the terms of a supply agreement between the organisation and the research tissue bank. TCCs are not research sites for the purposes of the RGF.

Research tissue bank managers are advised to provide R&D offices at all TCCs with a copy of the REC application for information, together with a copy of the favourable opinion letter when available. All TCCs should be listed in Part C of the REC application.

NHS researchers undertaking specific research projects using tissue or data supplied by a research tissue bank must apply for permission to R&D offices at all organisations where the research is conducted, whether or not the research tissue bank has ethical approval.

Site-specific assessment (SSA) is not a requirement for ethical review of research tissue banks.

Registration of Research Tissue Banks

It is a condition of the ethical approval that all Research Tissue Banks are registered on the UK Clinical Research Collaboration (UKCRC) Tissue Directory. The Research Tissue Bank should be registered no later than 6 weeks after the date of this favourable ethical opinion letter or 6 weeks after the Research Tissue Bank holds tissue with the intention to provide for research purposes. Please use the following link to register the Research Tissue Bank on the UKCRC Directory: <https://directory.biobankinguk.org/Register/Biobank> Registration is defined as having added details of the types of tissue samples held in the tissue bank.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity *e.g.*, when submitting an amendment or annual progress report form. We will monitor the registration details as part of the annual progress reporting process.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

Research tissue bank managers are advised to provide R&D offices at all TCCs with a copy of the REC application for information, together with a copy of the favourable opinion letter when available. All TCCs should be listed in Part C of the REC application.

NHS researchers undertaking specific research projects using tissue or data supplied by a research tissue bank must apply for permission to R&D offices at all organisations where the research is conducted, whether or not the research tissue bank has ethical approval.

Site-specific assessment (SSA) is not a requirement for ethical review of research tissue banks.

Registration of Research Tissue Banks

It is a condition of the ethical approval that all Research Tissue Banks are registered on the UK Clinical Research Collaboration (UKCRC) Tissue Directory. The Research Tissue Bank should be registered no later than 6 weeks after the date of this favourable ethical opinion letter or 6 weeks after the Research Tissue Bank holds tissue with the intention to provide for research purposes. Please use the following link to register the Research Tissue Bank on the UKCRC Directory: <https://directory.biobankinguk.org/Register/Biobank> Registration is defined as having added details of the types of tissue samples held in the tissue bank.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity *e.g.*, when submitting an amendment or annual progress report form. We will monitor the registration details as part of the annual progress reporting process.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached standard conditions give detailed guidance on reporting requirements for research tissue banks with a favourable opinion, including:

9. Notifying substantial amendments
10. Submitting Annual Progress reports

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:

<http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/>


HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at <http://www.hra.nhs.uk/hra-training/>

18/EE/0069

Please quote this number on all correspondence

Yours sincerely



Revd Dr Derek Fraser
Chair

E-mail: NRESCommittee.EastofEngland-CambridgeCentral@nhs.net

16 July 2018

Frank Martin / Ishaan Maitra
School of Pharmacy and Biomedical Sciences
University of Central Lancashire

Dear Frank / Ishaan

Re: STEMH Ethics Committee Application
Unique reference Number: STEMH 909

The STEMH ethics committee has granted approval of your proposal application 'Delineating the difficulties in diagnosing Oesophageal Cancer destined to arise from Barrett's Oesophagus from body fluid spectroscopy analysis'. Approval is granted up to the end of project date*. It is your responsibility to ensure that:

- the project is carried out in line with the information provided in the forms you have submitted
- you regularly re-consider the ethical issues that may be raised in generating and analysing your data
- any proposed amendments/changes - including transfer of samples to another researcher - to the project are raised with, and approved, initially by BTNW and then submitted to STEMH
- you notify EthicsInfo@uclan.ac.uk if the end date changes or the project does not start
- serious adverse events that occur from the project are reported to Committee
- a closure report is submitted to complete the ethics governance procedures (Existing paperwork can be used for these purposes e.g. funder's end of grant report; abstract for student award or NRES final report. If none of these are available use [e-Ethics Closure Report Proforma](#)).
- human tissue held under this project (which has been approved by BTNW) is stored and used in accordance with the HTA licence requirements. At the end of the project, any unused human tissue samples must be returned to BTNW for further use/storage or appropriate disposal. Samples that do not fall within the HTA's definition of 'relevant material' should be disposed of in accordance with all relevant H&S requirements including any specific BTNW disposal arrangements.

Yours sincerely,



Karen Rouse
Chair

STEMH Ethics Committee

cc UCLan HT Technician

* for research degree students this will be the final lapse date

NB - Ethical approval is contingent on any health and safety checklists having been completed, and necessary approvals as a result of gained.

10.5 Publications

1) Towards screening Barrett's Oesophagus: current guidelines, imaging modalities and future developments (Review article)

Maitra I, Date RS, Martin FL

Clin J Gastroenterol. 2020 Jun 3

PMID: 32495144 (**Chapter 1**)

2) Establishing spectrochemical changes in the natural history of oesophageal adenocarcinoma from tissue Raman mapping analysis

Maitra I, Morais CLM, Lima KMG, Ashton KM, Bury D, Date RS, Martin FL

Anal Bioanal Chem. 2020 Apr 25; 203(3):49-52.

PMID: 32333079 (**Chapter 7**)

3) Raman spectral discrimination in human liquid biopsies of oesophageal transformation to adenocarcinoma.

Maitra I, Morais CLM, Lima KMG, Ashton KM, Date RS, Martin FL

J Biophotonics. 2019 Dec 3; e201960132

PMID: 31794123 (**Chapter 4**)

4) Attenuated total reflection Fourier-transform infrared spectral discrimination in human bodily fluids of oesophageal transformation to adenocarcinoma.

Maitra I, Morais CLM, Lima KMG, Ashton KM, Date RS, Martin FL

Analyst. 2019 Dec 2;144(24):7447-7456.

PMID: 31696873 (**Chapter 3**)

5) Discrimination of oesophageal transformation stages to adenocarcinoma on human tissue samples using Raman microspectroscopy

Maitra I, Morais CLM, Lima KMG, Ashton KM, Bury D, Date RS, Martin FL

Vib Spectrosc. 2020 Sept 1; 203(3):49-52.

DOI: 10.1016/j.vibspec.2020.103141 (**Chapter 6**)

10.6 Presentations

1) Raman spectral discrimination in human liquid biopsies of oesophageal transformation to adenocarcinoma

Maitra I, Morais CLM, Lima KMG, Ashton KM, Date RS, Martin FL

55th Congress European Society for Surgical Research, Innsbruck, Austria

Accepted – presentation date 10th December 2020, **ORAL** (**Chapter 4**)

2) Attenuated total reflection Fourier-transform infrared spectral discrimination in human bodily fluids of oesophageal transformation to adenocarcinoma

Maitra I, Morais CLM, Lima KMG, Ashton KM, Date RS, Martin FL

55th Congress European Society for Surgical Research, Innsbruck, Austria

Accepted – presentation date 11th December 2020, **ORAL** (**Chapter 3**)

3) Raman spectral tissue discrimination in oesophageal transformation to adenocarcinoma

Maitra I, Morais CLM, Lima KMG, Ashton KM, Bury D, Date RS, Martin FL

55th Congress European Society for Surgical Research, Innsbruck, Austria

Accepted – presentation date 11th December 2020, **ORAL & POSTER** (**Chapter 6**)