Applications of Raman spectroscopy and fluorescence spectrometry in glioma surgery and neuropathology

Laurent James Livermore
MA(Cantab) MBBS MRCS



Kellogg College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy (D.Phil.)
in Clinical Neurosciences

Applications of Raman spectroscopy and fluorescence spectrometry in glioma surgery and neuropathology

Laurent James Livermore
Kellogg College, University of Oxford
A thesis submitted for the degree of Doctor of Philosophy (D.Phil.) in Clinical Neurosciences
Trinity term, 2019

Abstract

The diagnosis and management of patients with diffuse gliomas presents a significant challenge due to the aggressive nature of these tumours. Currently, there is no cure, and therefore the aim of treatment is to prolong survival and maintain quality of life for as long as possible. Surgical resection is an important part of the treatment, with evidence that increasing the extent of tumour resection results in prolonged survival, but the intraoperative identification of tumour boundaries is difficult. Techniques such as fluorescence-guided surgery have been shown to increase the extent of resection. The molecular genetic classification of gliomas, particularly identification of isocitrate dehydrogenase mutation and 1p/19q-codeletion, has also been shown to be critical for clinical and surgical decision-making.

Raman spectroscopy probes the unique molecular vibrations of a sample in order to accurately characterise its molecular composition. No sample processing is required, allowing for rapid analysis of tissue. The first aim of this project was to apply Raman spectroscopy to intraoperative and neuropathology analysis of gliomas both in terms of genetic subtyping and identification of the tumour/normal tissue boundary. *Ex vivo* Raman analysis of fresh tissue samples, from 62 patients collected straight from surgery, demonstrated 79%-94% sensitivity and 90-100% specificity for predicting the three most common glioma genetic subtypes. Raman spectroscopy was also compared with intraoperative visual fluorescence and was found to be significantly better at identifying tumour at the tumour boundary compared to the gold standard of visual fluorescence.

The second aim of the project was to develop a fluorescence spectrometry laboratory setup to allow for testing of enhanced fluorescence visualisation methods. Here, a laboratory-based fluorescence setup was correlated with the intraoperative light sources. A stable fluorescence standard was proposed and developed, using a laser dye, and was correlated with the intraoperative fluorescence. This experimental setup was used to prove the hypothesis that the operative endoscope is more sensitive at detecting visual fluorescence than the operative microscope.

Table of Contents

Abstract **2**

Acknowledgements **6**

Abbreviations **7**

List of publications **8**

Chapter 1. Introduction **10**

1.1 Diffuse gliomas terminology 10

1.2 Glioma biomarkers and integrated neuropathological diagnosis 12

1.3 Patient pathway 15

1.4 The role of surgical resection in diffuse gliomas 16

1.5 Intraoperative decision making and surgical adjuncts 21

1.6 Raman spectroscopy in neuro-oncology 27

1.7 Fluorescence in neuro-oncology 38

1.8 Project background 40

1.9 Project aims 41

1.10 Thesis outline 41

Chapter 2. Optimisation of sample preparation and Raman spectrometer configuration **42**

2.1 Introduction 42

2.2 Method 43

2.3 Results 48

2.4 Conclusions 53

Chapter 3. Experimental parameters and effect of laser exposure on Raman spectra **57**

3.1 Preliminary experiment: 4 second exposure versus 6 second exposure time 57

3.2 Investigation into the effect of laser exposure on surrounding tissue 62

3.3 Investigation into sample degradation over time 67

3.4 Allan variance 76

3.5 Conclusion.....82

Chapter 4. Rapid molecular genetic classification of gliomas using Raman spectroscopy 87

4.1 Introduction.....87

4.2 Methods.....87

4.3 Results.....102

4.4 Discussion 116

Chapter 5. Development of a laboratory-based fluorescence setup, fluorescence standard and comparison of operative microscope and endoscope fluorescence detection thresholds 122

5.1 Introduction.....122

5.2 General methods.....123

5.3 Characterisation of light sources 124

5.4 PpIX fluorescence and photochemical degradation 129

5.5 DCM laser dye 135

5.6 Investigating evidence of signal saturation 137

5.7 Correlation of PpIX and DCM fluorescence.....147

5.8 Comparison of microscope and endoscope fluorescence.....154

5.9 Conclusion.....162

Chapter 6. Identification of tumour versus normal brain using Raman spectroscopy ..164

6.1 Introduction.....164

6.2 Method 165

6.3 Results 169

6.4 Discussion 173

Chapter 7. Conclusions and future work 177

7.1 Optimisation of Raman spectroscopy experimental protocol 177

7.2 Rapid molecular genetic classification of gliomas using Raman spectroscopy 178

7.3 Development of laboratory-based fluorescence spectrometry setup 180

7.4 Comparing Raman spectroscopy and visual fluorescence 180

7.5 Closing remarks 183

Appendix A: Coating protocol 184

Appendix B: Fresh tissue 185

Appendix C: Peak Allocation 189

References 193

Acknowledgements

I would like to thank Cancer Research UK (CRUK) Oxford Centre for funding my Clinical Research Training Fellowship, giving me the opportunity to undertake this project.

I am indebted to many individuals for their help and support in the completion of this project and the writing of this thesis. First and foremost, I would like to thank my supervisors: Professor Claire Vallance, Professor Olaf Ansorge and Mr Puneet Plaha. Prof Vallance has been integral in guiding me through this project with her unparalleled efficiency and kindness. Her ability to explain complex chemistry and spectroscopic concepts in a simple fashion and her enthusiasm to pursue practical solutions to experimental challenges has been inspirational. Prof Ansorge has been an excellent academic mentor throughout this project. I have learnt from his amazing ability to generate scientific hypotheses, and propose methods to test them, which has helped me mature as a scientist. I am grateful for his constant support dealing with the challenges of developing fresh tissue neuropathological research within an NHS environment. Mr Plaha has been my inspiration for pursuing a career in neuro-oncology surgery. He has transformed the practice of neuro-oncology surgery in Oxford and is responsible for the initial idea to develop Raman applications in our neurosurgery group. I would like to thank him for facilitating my access to neurosurgery theatres throughout this project.

I am grateful to Renishaw Plc, our industry collaborator, specifically the Spectroscopy Products Division for giving me access to their Raman spectroscopy research facility and ultimately for the loan of a Raman spectrometer. I would particularly like to thank Dr Ian Bell and Dr Martin Isabelle. I have no doubt that this project would not have been a success without their ongoing commitment and support which far exceeded the expectations and time allotted to this project by the company. I would also like to acknowledge David Reece, Dr Stephanie Ohrel and Dr Katherine Lau who also helped to make this collaboration and project a success.

I would like to thank the Oxford Brain Bank staff, Carolyn Sloan and Hannah Brooks, for their help in accessing archived brain samples, and the John Radcliffe Hospital NHS Neuropathology Department for their support of the research. I am particularly indebted to Connor Scott for teaching me how to cut FFPE samples and for preparing the cryosections used in this project, and I would like to acknowledge Dr Joan Gannon and John Walsby-Tickle for providing the LN18 cell lines. I am grateful to the excellent John Radcliffe Hospital neurosurgery theatre scrub staff and to neuro-oncology surgeons, Mr Richard Stacey and Mr Vasileios Apostolopoulos, for helping me with intraoperative sample collection.

Finally, I would like to thank my wife, Rachel, for her unwavering support, patience and encouragement throughout this project, and my children, Sophie and Tom, who avoid disturbing me in the office when I'm working! I am also eternally indebted to my parents for everything they have done and provided for me.

Abbreviations

5ALA	5-aminolevulinic acid
ANOVA	Analysis of variance
ATR	Attenuated total reflectance
AUS	Area under spectrum
AUC	Area under curve
BBB	Blood-brain barrier
CaF_2	Calcium fluoride
CARS	Coherent anti-Stokes Raman scattering
CCD	Charge-coupled device
CNS	Central nervous system
CRS	Coherent Raman spectroscopy
DCM	4-Dicyanomethylene-2-methyl-6-p-dimethylaminostyryl-4H-pyran
DTI	Diffusion tensor imaging
EANO	European Association of Neuro-oncology
EMSC	Extended Multiplicative Scatter Correction
EOR	Extent of resection
EtOH	Ethanol
FFPE	Formalin-fixed paraffin-embedded
FISH	Fluorescence in-situ hybridization
FLAIR	Fluid attenuated inversion recovery
FTIR	Fourier transform infrared spectroscopy
GB	Glioblastoma
GTR	Gross total resection
HGG	High grade glioma
IDH	Isocitrate dehydrogenase
LED	Light emitting diode
LDA	Linear discriminant analysis
LGG	Low grade glioma
MDT	Multidisciplinary team
MRI	Magnetic resonance imaging
MUT	Mutated
NICE	National Institute for Health and Care Excellence
OS	Overall survival
PC	Principal component
PCA	Principal component analysis
PFS	Progression free survival
PpIX	Porphyrin IX
RCT	Randomised control trials
SiO2	Fused silica
SNR	Signal-to-noise ratio
SRS	Stimulated Raman scattering
SS	Stainless steel
WT	Wildtype
WHO	World Health Organisation

List of publications

Publications arising from this thesis at time of submission:

Livermore LJ, Isabelle M, Bell IM, Scott C,Walsby-Tickle J, Gannon J, Plaha P, Vallance P, Ansorge O. Rapid intraoperative molecular genetic classification of gliomas using Raman spectroscopy. ***Neuro-oncology Advances***. (2019). In Press

Livermore LJ, Isabelle M, Bell I, Plaha P, Vallance C, Ansorge O. Rapid Genetic classification of gliomas Using Raman spectroscopy. (Abstract) **Neuro-oncology.** 2018;20(suppl_5): 346-v.

Goodacre R, Sergo V, Barr H, Sammon C, Schultz ZD, Baker MJ, Graham D, Marques MP, Sulé-Suso J, **Livermore J**, Faulds K, Sinjab F, Matousek P, Campbell CJ, Dluhy R, Gardner P, Phillips C, Diem M, Wood B, Apolonskiy A, Kazarian S, Fullwood L, Gough K, Petrich W, Lloyd G, Ibrahim O, Cinque G, Sockalingum GD, Stone N, Kendall C, McAughtrie S, Perez-Guaita D, Clark L, Gerwert K, Bonifacio A, Notingher I, Lasch P, Bhargava R, Lepert G, Mader K, Paterson C. Clinical Spectroscopy: general discussion. ***Faraday Discuss.*** 2016 Jun 23; 187:429-60

Presentations arising from this thesis at time of submission:

Invited speaker

SPEC 2018, International Clinical Spectroscopy Conference, Glasgow, UK (2018)
Hot topic discussion panel – How do we translate this to the clinic?

CLIRCON17 - The International Society for Clinical Spectroscopy, UK (2017)
Raman spectroscopy for the identification of isocitrate dehydrogenase (IDH) mutated glioblastomas (Invited speaker) ***LJ Livermore***

SPEC 2016, International Clinical Spectroscopy Conference, Montreal, Canada (2016)
Translation to clinic, discussion group

Oral presentations

SPEC 2018, International Clinical Spectroscopy Conference, Glasgow, UK (2018)
Rapid genetic classification of gliomas using Raman spectroscopy.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

Society of British Neurological Surgeons, Spring Meeting 2017, Oxford, UK (2017)
Raman spectroscopy for the identification of isocitrate dehydrogenase (IDH) mutated glioblastomas.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

European Association of Neurosurgical Societies (EANS) Congress 2016, Greece (2016)
Raman spectroscopy for the identification of isocitrate dehydrogenase (IDH) mutated glioblastomas.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

SPEC 2016, International Clinical Spectroscopy Conference, Montreal Canada (2016)
Raman spectroscopy for the identification of isocitrate dehydrogenase (IDH) mutated glioblastomas.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

British Neuro-Oncology Society Conference 2016, Leeds, UK (2016)
Raman spectroscopy for the identification of isocitrate dehydrogenase (IDH) mutated glioblastomas.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

Poster presentations

CRUK brain tumour conference, London, UK (2018)
Rapid genetic subclassification of gliomas using Raman spectroscopy.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

CLIRCON17 - The International Society for Clinical Spectroscopy, UK (2017)
Allan Variance - a useful method of establishing the optimum laser exposure for Raman studies of tissue samples.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

European Association of Neuro-oncology (EANO) Annual meeting, Germany (2016)
Raman spectroscopy for the identification of isocitrate dehydrogenase (IDH) mutated glioblastomas.
L. J. Livermore, I. Bell, M. Isabelle, C. Scott, P. Plaha, C. Vallance, O. Ansorge

Chapter 1. Introduction

The field of neuro-oncology surgery has transformed over the last twenty years. Patients with aggressive cancerous brain tumours were once operated on by junior neurosurgical registrars with the aim of surgery being to "give the brain a bit of space and *maybe* the patient a bit of time". Now the role and benefit of neurosurgical resection in the management of these tumours has been more clearly defined, and its practice has become a highly specialised branch of neurosurgery employing advanced surgical techniques. A significant catalyst to this change has been advances in the neuropathological diagnosis of these tumours, which has moved from the standard histological evaluation to targeted molecular and genetic tumour fingerprinting that defines not only tumour subtypes but also dictates treatment decisions. Despite these advances, the diagnosis of a malignant brain tumour is devastating for patients and their relatives, with a median survival of 14 months even with full medical and surgical treatment[2]. Brain tumours remain a cancer with substantial unmet need, accounting for more average years of life lost than all other cancers[3,4].There is no cure currently for these tumours, and therefore the aim of treatment is to prolong survival and maintain quality of life for as long as possible. This balance between quality and quantity of life is essential given the short overall gain in length of life associated with any treatment. This thesis reports on the development and application of spectroscopic techniques, specifically Raman spectroscopy and fluorescence spectrometry, in both neurosurgery and neuropathology, with the aim of improving the diagnosis and treatment of patients with these tumours.

1.1 Diffuse gliomas terminology

Gliomas are primary brain or spine tumours which originate from glial cells. The most common of these in the adult population are tumours arising from astrocytes or their precursors, known as astrocytomas, and tumours arising from oligodendrocytes or their precursors, known as oligodendrogliomas. The World Health Organisation (WHO) Classification of Tumours of the Central Nervous System (CNS)[5,6i] categorises over 130 different types of brain tumour, and within this there are several different groups of gliomas. The largest group is referred to as "Diffuse astrocytic and oligodendroglial tumours". This is

[i] The *WHO Classification of Tumours of the Central Nervous System* is the recognised international standard for classification of central nervous system tumours.

also the most common group of tumours seen in clinical practice, accounting for nearly a third of all primary brain tumours. They are also the most common malignant brain tumour[7]. Other glial tumours, such as pilocytic astrocytomas, diffuse midline gliomas (also in the aforementioned group) or ependymomas are rarer entities that are not covered here. Thus, when referring to glioma in this thesis, this refers to astrocytomas and oligodendrogliomas.

The histological diagnosis of CNS tumours includes an assessment of tumour grade (WHO grade). The WHO grading for CNS tumours is from one to four (WHO I to IV) with increasing grades corresponding to increased malignancy and growth rate of the tumour. WHO I tumours are generally benign and slow growing. WHO II tumours are also slow growing but, in the context of astrocytomas and oligodendrogliomas, they will normally progress to a higher grade over time. WHO III tumours are considered malignant and grow fast, and WHO IV tumours have the highest level of malignancy and growth. By definition, astrocytomas can only be assigned WHO II, III or IV and oligodendrogliomas can be assigned either WHO II or III. The term "low grade glioma" (LGG) is used to refer to WHO II astrocytomas or oligodendrogliomas, while "high grade glioma" (HGG) refers to WHO III or IV gliomas[ii]. In the WHO classification WHO II gliomas are labelled as "diffuse", WHO III as "anaplastic" and WHO IV astrocytomas are labelled as "glioblastomas". These somewhat confusing labels are summarised in Table 1.1.

The natural history of both astroglial and oligodendroglial low grade gliomas is that they will eventually transform to high grade tumours. Glioblastomas that arise due to transformation of a low grade glioma into a WHO III and then WHO IV tumour are historically referred to as "secondary glioblastomas" as they originate from a lower grade tumour. Glioblastomas can also arise *de novo* without transitioning through the lower grades and are historically known as "primary glioblastoma". There is a clear difference in the cascade of genetic mutations that occur in primary compared to secondary

[ii] The terms "low grade glioma" and "high grade glioma" are historical terms that are not featured in the current WHO classification. They are still, however, commonly used in both clinical and research settings and thus are defined and used here.

glioblastomas, and this forms an important part of the recent reclassification of these tumours outlined in the next section.

	WHO grade	Histological type	
		Astrocytoma	**Oligodendroglioma**
"Low grade glioma"	WHO II	Diffuse astrocytoma	Diffuse oligodendroglioma
"High grade glioma"	WHO III	Anaplastic astrocytoma	Anaplastic oligodendroglioma
	WHO IV	Glioblastoma	

Table 1.1 - Summary of glioma terminology regarding WHO grade and histological type.

1.2 Glioma biomarkers and integrated neuropathological diagnosis

The WHO Classification of CNS tumours was revised in 2016 to incorporate molecular genetic parameters along with histological, phenotypic parameters to establish an "integrated" neuropathological diagnosis[5,8]. This reflects the increasing understanding of the genetic basis of tumourigenesis and predictive and prognostic consequences of certain genetic subtype. In terms of diffuse astrocytomas and oligodendrogliomas the key mutations are mutations of the isocitrate dehydrogenase (IDH) genes 1 or 2 and chromosome 1p and 19q (1p/19q) codeletion.

1.2.1 Isocitrate dehydrogenase (IDH) mutations

The discovery of IDH mutations in most WHO grade II and III gliomas and secondary glioblastomas was a ground-breaking advance in the understanding of this disease[9-11]. IDH enzymes normally catalyse the decarboxylation of isocitrate to generate α-ketoglutarate (α-KG). A mutation in either IDH1 or IDH2 confers a neomorphic enzyme activity that catalyses reduction of α-KG to D-2-hydroxyglutarate (2HG). This inhibits α-KG-dependent dioxygenases, which results in aberrant DNA and histone methylation, eventually leading to widespread hypermethylation of CpG islands within the cell DNA[1,12,13]. IDH1 and IDH2 mutations are similar in structure and have identical reactions; thus mutations in either gene have a similar downstream effect. The most common mutation is a single base transition substitution of arginine for histidine in IDH1, denoted IDH1(R132H), which accounts for nearly 90% of the mutations[13].

Within the laboratory setting, IDH mutation is detected using immunohistochemistry (IHC), with a specific monoclonal antibody for the common IDH1(R132H) mutation[14]. This will give a false negative result if a rare IDH mutation is present within the tumour, and thus, if there is clinical or histological suspicion of a rare IDH mutation then the sample will undergo targeted IDH1/IDH2 genetic sequencing. The IHC test takes around 24 hours to complete.

As well as being an integral part of the neuropathological classification of glioma, IDH mutations also have clinical significance. The presence of an IDH mutation is a strong predictor of longer progression free survival[15-17] and overall survival[9,10,15-22] compared to IDH-wildtype tumours. Indeed, some studies have shown that IDH mutation status may be more important than histological grade in predicting prognosis[23]. There is also some evidence that IDH-mutant tumours respond better to radiochemotherapy, although the evidence here is somewhat conflicting[24-26].

1.2.2 1p/19q-codeletion

1p/19q-codeletion is the hallmark of oligodendroglioma. The 2016 revised WHO classification stipulates that 1p/19q-codeletion and IDH mutation must be present in order to make the diagnosis of oligodendroglioma. Within most neuropathology laboratories, 1p/19q-codeletion is routinely established using the cytogenetic technique of fluorescence in-situ hybridization (FISH). 1p/19q-codeletion is a strong favourable prognostic factor for overall survival[17,20,27] and a strong predictor of response to radiochemotherapy[28-31].

1.2.3 Other diagnostic and prognostic markers

In addition to IDH and 1p/19q-codeletion, other biomarkers have been investigated with a view to targeting treatments, predicting responses to therapy, and improving prognostication. The most clinically relevant of these is MGMT (O^6-methylguanine-DNA methyltransferase) promoter hypermethylation, which is a strong predictor of glioblastoma response to alkylating chemotherapy agents such as temozolomide[2,32,33] and is often used to determine the use of such agents in elderly patients with glioblastoma[34].

In terms of establishing an integrated neuropathological diagnosis, other important markers include ATRX expression, TERT-promoter mutations, KIAA1549–BRAF fusions, BRAF-V600E mutation, and H3F3A-G34 mutation. ATRX (alpha-thalassemia/mental retardation syndrome X-linked) expression is

particularly important in generating an integrated diagnosis of diffuse glioma. ATRX mutation leads to an IHC detectable loss of nuclear expression of the transcriptional regulator ATRX, which has an important role in chromatin remodelling and regulation of telomere length[35]. Loss of ATRX is associated with IDH-mutant astrocytomas but not IDH-mutant oligodendrogliomas, allowing for another level of differentiation in the neuropathological diagnostic algorithm of glioma (Figure 1.1).

1.2.4 Integrated neuropathological diagnosis

The combination of the aforementioned biomarkers along with histological evaluation allows the neuropathologist to generate an integrated histological and molecular classifiation of the diffuse gliomas. A simplified summary of the diagnostic algorithm for astrocytomas and oligodendrogliomas is shown in Figure 1.1, and results in five diagnostic classes:

1) glioblastoma, IDH-mutant (WHO IV);
2) diffuse or anaplastic astrocytoma, IDH-wildtype (WHO II or III);
3) glioblastoma, IDH-wildtype (WHO IV);
4) diffuse or anaplastic astrocytoma, IDH-mutant (WHO II or III);
5) diffuse or anaplastic oligodendroglioma, IDH-mutant and 1p/19q-codeleted (WHO II or III).

IDH-wildtype WHO II or III astrocytoma is considered a provisional entity in the 2016 WHO Classification as these are most likely undersampled glioblastoma[36].

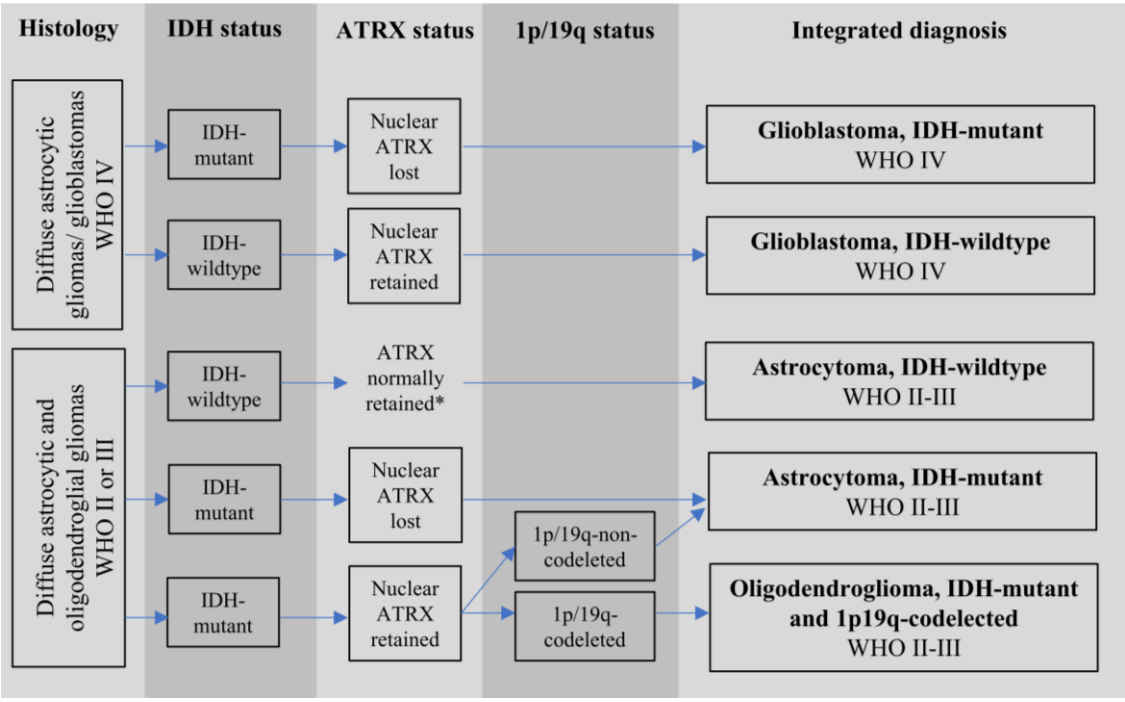


*Figure 1.1 - Simplified diagnostic algorithm for neuropathological integrated diagnosis of astrocytoma and oligodendroglioma (modified from Reifenberger et al.[1]). *ATRX expression is retained in most IDH-wildtype WHO II-III astrocytomas.*

1.3 Patient pathway

The majority of patients with a brain tumour present as an emergency[37] and are initially managed by non-specialists centres who seek advice from the neuro-oncology multidisciplinary team (MDT[iii]). The MDT meeting acts as a focus for the management of these patients. Patients will routinely have a (gadolinium) contrast-enhanced magnetic resonance imaging (MRI) scan of the brain, and possibly a whole body CT (computed tomography) scan if indicated to look for another source of primary cancer (the most common adult brain tumour overall is a metastasis). Imaging normally provides sufficient information to narrow the differential diagnosis down to a few possibilities. If diffuse low grade or high grade glioma is suspected then, depending on the patient's age, clinical condition, and performance status[iv], surgery may be offered. In most cases this is in the form of surgical resection of the tumour. If surgical resection is not appropriate or is declined by the patient, a biopsy of the tumour is often undertaken in order to establish the neuropathological diagnosis. After surgery the patient is discussed again at the MDT

[iii] The neuro-oncology MDT normally consists of neurosurgeons, neuro-oncologists, neuroradiologists, neuropathologists and neuro-oncology specialist nurses. Specialist radiographers, palliative care specialists, neuropsychologists, physiotherapists or occupational therapists may also form part of the MDT in some units.
[iv] Performance status is a measure of a patient's ability to undertake activities of daily living and self-care. Commonly used scoring systems include Karnofsky scoring and the WHO performance status.

meeting, with the postoperative MRI showing the extent of resection and the formal neuropathological diagnosis. At this stage the adjuvant treatment course is defined, and would normally consist of the options of active monitoring with regular MRI scans, chemotherapy alone, radiotherapy alone, or a combination of chemotherapy and radiotherapy. Guidelines for treatment are published by the European Association of Neuro-oncology (EANO)[38] and also by the National Institute for Health and Care Excellence (NICE) who have recently published a new set of guidelines for the management of adult glioma and brain metastasis in 2018[34,39,40].

1.4 The role of surgical resection in diffuse gliomas

1.4.1 Aim of surgery

The suitability of a patient to undergo surgical reaction for their brain tumour depends on many factors, including age, performance status, and the patient's wishes. The location of the tumour is also important, as many parts of the brain are known to have important function (known as eloquent brain) and cannot be resected without high risk of severe or devastating disability or risk to the patient's life. If surgical resection is decided to be the best course of action, then the aims of surgery are: (1) to take tissue biopsies for neuropathological diagnosis and plan postoperative treatment; and (2) to remove as much of the tumour as safely possible. This second aim, also known as "maximum cytoreduction", is the main consideration of the neurosurgeon when planning and undertaking glioma surgery, but this is constantly being balanced against the risk of neurological disability (Figure 1.2).

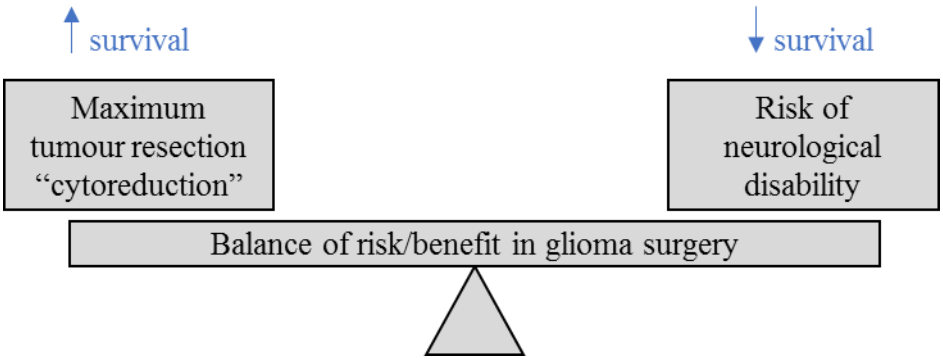


Figure 1.2 - Surgical decision making in glioma resections. The balance between survival benefit of maximum tumour resection and the risk of neurological disability which decreases survival.

The driver for achieving maximal tumour resection is the principle that the more tumour that is removed, i.e. the extent of resection (EOR), the greater the effect on progression free survival, overall survival, and symptomatic relief. Complete resection is, however, not possible. Diffuse gliomas, as their name

suggests, infiltrate throughout the brain[41] and do not grow as a well-defined mass with a clear interface between tumour and normal brain. Even if a surgical plane is seen between tumour and brain, there will always be microscopic penetration of tumour cells that cannot be completely resected. This gradual transition from solid tumour to normal brain is often seen in histological section, and is known as the transition zone. 80% of glioblastoma recurrences occur within the tumour resection cavity[42], supporting the fact that microscopic complete resection is rarely achieved. The function of the surrounding brain is also important to consider: resecting beyond the visual tumour boundary may not be safe given the risk of removing eloquent brain and causing neurological disability. The diffuse nature of the tumours and the importance of the surrounding brain tissue means that the concept of a resection margin, where a rim of normal tissue is included in the resection as might be done in lung, liver or skin cancer surgery, cannot be applied to glioma surgery.

1.4.2 Defining extent of resection (EOR)

The extent of surgical resection is assessed based on comparison between preoperative and postoperative MRI scans. In HGG, where contrast agent is typically taken up by the tumour, the aim of surgery is to remove the contrast-enhancing component as assessed on postoperative T1-with-contrast MRI scan. LGG tend not to enhance on MRI and are typically best viewed on a T2/fluid attenuated inversion recovery (FLAIR) sequence. The aim of surgery is to achieve complete resection of the FLAIR signal, as assessed on postoperative MRI. It is well recognised that both FLAIR and T1 contrast enhancement will underestimate the extent of tumour[43,44]. If postoperative MRI confirms complete resection of the contrast-enhanced or FLAIR regions then the term "gross total resection" (GTR) is used, acknowledging that there will always be some tumour left behind. If there is residual tumour on postoperative imaging, then the resection is referred to as "subtotal". The concept of "near-total" resection is also used, normally referring to 90 or 95% EOR.

1.4.3 Effect of EOR on survival

There is no class I evidence to support the hypothesis that increased EOR has a beneficial effect on progression free survival (PFS) or overall survival (OS), and a trial randomising patients into different planned EOR groups is unlikely to pass an ethics committee. The body of evidence has mainly come from several retrospective studies reviewing the EOR and corresponding PFS and/or OS. There is also

some prospective data extrapolated from randomised control trials (RCTs) investigating the effect of surgical techniques on EOR and survival. A criticism of these studies is the influence of selection bias based on other factors, such as tumour size, eloquent location influencing the ability to achieve EOR, and thus the causative effect of EOR. Modern MRI viewing software allows for accurate volumetric assessment of tumour, and thus recent studies report quantitative measures of EOR and residual tumour volumes. This technology was not available for earlier studies, and thus more qualitative, non-volumetric, measures of EOR were reported based on assessment by neuroradiologists.

With regards to HGG, some early reports in the 1990s using nonvolumetric imaging analyses of EOR failed to show clear benefits of increasing EOR on survival[45-47]. These have mostly been superseded by large series using both nonvolumetric[48-62] as well as detailed volumetric[63-71] imaging analyses of EOR, which support the hypothesis that increased EOR is a predictor of both improved PFS and OS. Volumetric studies have also reported a stepwise improvement in survival benefit with increased EOR, showing that achieving at least 70% resection provides survival benefit in HGG, with incremental benefit in survival continuing even at the highest levels of resection[68,69,72]. Evidence supporting the benefit of EOR has also been extrapolated from an RCT designed to evaluate the benefit of 5-aminolevulinic acid-induced fluorescence-guided surgery[73] and another RCT evaluating intraoperative MRI[74]. Both studies demonstrated increased rate of gross total resection with the use of these techniques and a corresponding increase in 6-month PFS. Finally, a recent meta-analysis of the association of EOR with survival in glioblastomas reviewed 37 studies and found that gross total resection was significantly associated with a decreased likelihood of disease progression at 6 months and 1 year and a lower relative risk of mortality at 1 and 2 years compared to subtotal resection[75].

The body of evidence for EOR in LGG is of higher quality, given the previously recognised treatment option of active monitoring with regular MRI scans in these patients. This has allowed for a number of prospective cohort studies evaluating a range of treatment options, from diagnostic biopsy and active monitoring to gross total resection[76]. In both WHO II astrocytomas and oligodendrogliomas, studies using nonvolumetric imaging analyses[48,77-91] and volumetric analyses[70,92-95] have shown that greater EOR as measured by T2/FLAIR signal is associated with better PFS and OS. The time to malignant transformation is also increased with greater EOR[94-96]. This has led to a paradigm shift in the management

of these tumours, away from a biopsy and active monitoring for several years and instead offering the patient the option of early surgical resection at the time of diagnosis.

One important consideration when interpreting the above EOR data is that the studies were carried out prior to the current era of defining glioma types using molecular genetic markers which have known predictive and prognostic values. These studies therefore do not take these important molecular characteristics, and possible confounding factors for survival outcomes, into account. To this end, more recent studies have retrospectively looked at the effect of EOR on survival outcomes in the context of IDH and 1p/19q-codeletion.

In high grade astrocytoma (anaplastic astrocytoma (WHO III) and glioblastoma(WHO IV)), Beiko *et al.* found that the presence of IDH mutation was an independent predictor of gross total resection, compared to IDH-wildtype, regardless of patient age, tumour grade or location[97]. They also report an increased survival benefit for gross total resection in IDH-mutant compared to IDH-wildtype HGG. Kawaguchi *et al.* report the impact of EOR on WHO III gliomas, comparing oligodendroglioma (1p/19q-codeleted, IDH-mutant), IDH-mutant astrocytoma and IDH-wildtype astrocytoma[98]. They found a significant survival advantage between achieving gross total resection over subtotal resection in IDH-mutant astrocytoma but no significant survival advantage for oligodendroglioma or IDH-wildtype astrocytoma. In WHO II gliomas, Cardier *et al.* report that the ability to achieve total resection was independent of IDH mutation or 1p/19q-codeletion[99]. However, this study did not aim to correlate their findings with survival outcomes. To date, the largest and most robust volumetric assessment of EOR and survival on genetically stratified WHO II gliomas was carried out by Wijnenga *et al.*[100]. This study, after correcting for molecular subtypes, confirms the conclusion of the pre-genetic stratification studies that greater EOR is associated with longer OS. They do however find a difference between the subtypes in terms of response to EOR, reporting that the effect of less tumour residual volume on overall survival was much greater in IDH-mutant astrocytomas compared to oligodendrogliomas. In the IDH-mutant astrocytoma group, even a small amount of residual tumour of 0.1 cm^3 impacted negatively on survival compared to 0 cm^3 residual. This was not seen in the oligodendroglioma group, with a small residual having no significant effect on overall survival.

In general, these studies suggest that the survival benefit of EOR is partly dependent on the molecular stratification of the tumour. When the surgeon is resecting at the boundary of the tumour and the risk of disability from further surgical resection is high, information regarding the genetic diagnosis of the tumour could strongly influence decision making. If the tumour is found to be an IDH-mutant astrocytoma (either WHO II, III or IV (glioblastoma)), the balance of risk/benefit might result in pushing the resection to its maximum. If, however, the tumour is found to be an oligodendroglioma or IDH-wildtype astrocytoma of any grade, there may not be benefit in risking neurological function to achieve a maximal resection. This makes a strong argument for obtaining molecular-level diagnostic information on gliomas at the time of surgical resection, which is the main motivation for the application of Raman spectroscopy explored in this thesis.

1.4.4 Beyond gross total resection – functional guided resection

It is now accepted in both high and low grade gliomas that every effort should be made to achieve the maximum surgical resection of the tumour that is safely possible. Furthermore, it is proposed by some that there should be an attempt to go beyond the recognised tumour margins as defined by preoperative MRI and extend the resection as far as the surrounding functional brain will allow. If there is no eloquent function in the brain tissue surrounding a tumour, and it is generally accepted that there will be tumour cells within this region, then it is logical to suggest that extending resection beyond the FLAIR or contrast enhancing signal would possibly be of benefit. In LGG there is an established school of thought, championed by Hugues Duffau[96,101,102], that the limit of the resection should be dictated by function only. In HGG this concept has not been universally adopted. Some retrospective series have looked at the effect on survival of extending the resection beyond the contrast enhancing region, and have shown that there is an increase in PFS and OS [71,103]. A non-randomised prospective study undertaken by our group showed a significant increase in PFS between patients who underwent functional guided resection beyond gross total resection versus gross total resection, independent of other prognostic factors[104], giving further evidence that this approach might be appropriate in selected HGG operations.

1.4.5 Postoperative neurological deficit and performance status

A permanent neurological deficit after resection surgery must be avoided for several reasons. Permanent postoperative neurological deficit and poor performance status are clear predictors of overall survival in

both low grade and high grade gliomas[66,95,105]. Adjuvant therapies will only be offered if the patient has a good performance status. In the case of glioblastomas, the 2018 NICE guidelines recommend the use of radiotherapy and concurrent chemotherapy only if the patient is able to be totally self-caring even if unable to carry on normal activity or work (Karnofsky Performance Score of ≥70)[34]. Leaving a patient with a dense hemiparesis after surgical resection would mean they would not meet this criterion.

Another important consideration is quality of life. Surgery does not cure patients of either low grade or high grade gliomas, but rather is a life-extending procedure, and therefore quality as well as quantity of life is an important consideration for patients and their clinicians. In the case of a newly diagnosed glioblastoma, the survival advantage of full treatment (surgery and adjuvant chemoradiotherapy) is likely to be in the region of only 12 months[2], and it is therefore important that patients make a value-based decision regarding treatment choices given the risk of neurological deficit.

1.5 Intraoperative decision making and surgical adjuncts

Several processes and technologies are available to help minimise the risk to the patient and to maximise safe surgical resection. Listed under the surgical question they help to answer, they include:

Question: What is the tumour type?

- intraoperative neuropathology

Question: Is there tumour present?

- neuronavigation
- intraoperative MRI (iMRI)
- intraoperative ultrasound (iUS)
- fluorescence-guidance surgery

Question: Is there eloquent brain which would result in neurological deficit if removed?

- Functional MRI
- Diffusion tensor imaging (DTI)
- awake surgery with cortical and subcortical stimulation

1.5.1 What is the tumour type? Intraoperative neuropathology

Whilst resecting the tumour the surgeon can send a sample to neuropathology and request an intraoperative histological analysis in order to obtain some information about the tumour at the time of surgery. This is particularly useful if the preoperative investigations have provided a broad range of differential diagnoses. Most commonly, the neuropathologist employs a smear preparation, in which a small tissue sample is smeared between two glass slides in order to quickly create a thin layer of section. This is then stained with haematoxylin and eosin (H&E) and reviewed under the microscope by the neuropathologist, taking about 30 minutes to complete. Information from smear samples is normally limited to the likely cell type of origin (differentiating between glioma, metastasis, and lymphoma, for example) and an idea of tumour grade. This information is used by the surgeon and can alter the surgical plan, especially if the tumour turns out to be different from what was thought from preoperative clinical assessment and imaging. An example in which the smear has a profound effect on surgical strategy would be the unexpected finding of a lymphoma. In this case the operation would normally be terminated after limited debulking of the tumour, as there are good non-surgical treatments for these tumours.

1.5.2 Is there tumour present? Surgical adjuncts to facilitate tumour resection

A recent Cochrane systematic review evaluated the evidence for using intraoperative imaging technology (neuronavigation, iMRI, iUS, fluorescence-guided surgery) to maximise extent of resection of gliomas[106]. The reviewers identified four randomised trials (two for iMRI, one for fluorescence-guided surgery and one for neuronavigation, and none for iUS) which all focused on HGG. Both iMRI trials were reported to provide very low-quality evidence, the fluorescence-guidance reported as low-quality, and the neuronavigation trial did not provide enough data to evaluate the effect on EOR. Citing multiple biases in these studies, the authors concluded that iMRI and fluorescence-guidance may be of benefit in maximising EOR, but that the quality of the evidence is very low. Despite this, recent NICE guidelines state that fluorescence-guidance and neuronavigation should be offered to all patients if appropriate, and iMRI and/or iUS should be considered if available in that centre[34].

Image guidance

Image guidance, or neuronavigation, systems allow the surgeon to estimate how far they have resected in relation to the preoperative MRI. This is used mainly to identify the contrast enhancing or T2/FLAIR

signal boundary in HGG and LGG respectively. Use of neuronavigation systems has become routine practice in most cranial neurosurgical procedures, including tumour surgeries, and is used to plan the skin incision and craniotomy as well as for intraoperative guidance within the brain tissue. Despite its routine use, there is no clear benefit in terms of EOR in glioma surgery[106]. The main shortcoming of neuronavigation is the effect of brain shift due to tumour resection or release of cerebrospinal fluid (CSF), which decreases the accuracy of the correlation with the preoperative imaging.

Intraoperative MRI and ultrasound

iMRI allows intraoperative correction of the brain shift, making neuronavigation more accurate, as well as identification of residual tumour, and has been shown to increase EOR and OS in patients with glioma[107-111]. The largest randomised trial comparing iMRI-guided with conventional image guidance surgery for HGG demonstrated a higher rate of complete resection in the iMRI group (98%) compared to the control group (68%), but this was not associated with an increase in PFS between the two groups[74]. There are several shortcomings to iMRI in its present state of development. Firstly, iMRI gives a single snapshot at a certain point in the resection process, rather than providing continuous feedback to the surgeon. This requires a careful choice of timing for the scan in relation to the surgical resection in order to achieve maximum benefit from the scan information. Secondly, aquiring the scan requires a significant pause in surgery, with the increase in overall surgical time reported to be between one and two hours[74,109]. Another major obstacle to the widespread use of iMRI is the significant installation cost of an integrated MRI operating theatre.

iUS, in contrast to iMRI, is relatively inexpensive, and can be used as many times as needed during an operation[112,113]. iUS can be integrated with the neuronavigation systems, and has been shown to be comparable in efficacy to iMRI for detecting tumours and guiding excisions[114]. However, the success of the technique is user dependent, and the method has low specificity for tumour. To date there are no studies investigating the effect of using iUS on the EOR in glioma surgery.

Fluorescence-guided surgery – Sodium fluorescein and 5-aminolevulinic acid

A number of fluorophores have been investigated for fluorescence-guided surgery, but only sodium fluorescein (NaF) and 5-aminolevulinic acid (5ALA) induced fluorescence have shown any potential in neurosurgery.

Sodium fluorescein

Sodium fluorescein is a small organic molecular salt with a blue excitation peak in the region of 465-490 nm and a green emission peak at 510 – 530 nm. The sensitivity of fluorescein to the presence of glioma has been reported at around 85% - 100%[115-117]. Several small, non-randomised, studies have shown the use of NaF to increase the rate of GTR[115,118,119] but there is concern about the low specificity of NaF, as the compound accumulates in areas of blood-brain barrier (BBB) breakdown or compromise[120]. Accumulation in regions of BBB compromise often includes fluorescence of peritumour oedematous regions, which could result in the surgeon removing potentially normal brain[121,122]. For this reason, and also due to the increasing popularity of fluorescence-guided surgery using 5ALA, the interest in NaF is currently waning.

5ALA-induced fluorescence

5ALA is naturally present in the human body as part of the haem synthesis pathway. Administration of exogenous 5ALA overloads the haem synthesis pathway, which causes accumulation of porphyrin-IX (PpIX), the last molecule prior to haem in the pathway[123]. Excitation of PpIX by 375-440 nm blue light causes emission of fluorescence in the region 630-710 nm[124-127]. Importantly, PpIX accumulation is selective for tumour, particularly HGG, with only low concentrations accumulating in normal brain[128-134]. The exact reason for high accumulation within tumour cells is still unclear, and this is an area of active research[123]. There is also evidence of selective accumulation in LGG, metastases, meningiomas and ependymomas[135-138]. In HGG the fluorescence correlates very well with gadolinium enhancement on preoperative MRI[131,134], and some have shown PpIX fluorescence to extend beyond, and to be more accurate than, contrast enhancement[139,140] with regards to determining the extent of tumour infiltration.

The popularity of 5ALA-induced fluorescence-guided surgery began with the publication of a landmark paper by Stummer in 2006[73]. This randomised controlled multicentre phase 3 trial compared fluorescence-guided resection using 5ALA with standard "white light" microsurgical resection, and demonstrated gross total resection in 65% of cases using 5ALA compared with 36% of cases in the white light group. This correlated with a significant increase in progression free survival and, on later view of patients where GTR was achieved, increased OS[141]. Despite criticism of this trial, including patient age and performance status and tumour location selection bias[106], 5ALA use has been adopted by many and

is recommended for the resection of HGG in the 2018 NICE guidelines[34]. Since this trial there have also been several case series reporting the use of 5ALA, often in combination with functional mapping, reporting resection rates much higher than the 65% achieved in the original study[140,142-145].

5ALA is administered as a drink dissolved in water, typically at a dose of 20 mg/kg[121,146]. Peak fluorescence occurs at around 6 to 8 hours after ingestion of 5ALA-in humans[146], so oral administration is normally around 4 hours prior to induction of anaesthesia so that peak fluorescence coincides with resection of the tumour. All the leading manufacturers of neurosurgery operative microscopes have the option of compatibility with 5ALA-induced fluorescence-guided surgery, which requires a blue light source for PpIX excitation and objective filters to capture only the emitted fluorescence. Using such instruments, the surgeon can switch between white and blue light illumination with minimal interruption to the surgical workflow. It is this integration into the normal surgical workflow that makes the technique so appealing to surgeons. and one reason why 5ALA has been taken up to such a great extent.

There is good correlation between the presence of visual PpIX fluorescence and the presence of HGG, giving 5ALA-induced fluorescence a high positive predictive value. It is known, however, that the negative predictive value of the absence of fluorescence is much lower, with up to 70% of biopsies taken beyond the visual fluorescence boundary still containing tumour[131,134]. The use of intraoperative fluorescence probes has been reported to detect fluorescent tumour beyond the visual fluorescence boundary. This will be explored further in Section 1.7.

1.5.3 Is there function? Surgical techniques to establish the presence of eloquent brain

As well as establishing the presence of tumour and normal brain at the boundary of resection, it is also important to establish the presence of brain function that, if the tissue were to be resected, could result in postoperative neurological deficit.

Functional MRI and diffusion tractography imaging

Functional MRI (fMRI) allows for cortical grey matter functions to be identified preoperatively and therefore for eloquent regions of the brain to be avoided during surgery. The lack of access to fMRI is the most limiting factor for its routine use. DTI is a method of estimating the location of the key white matter tracts within the brain and around the tumour. Tracts are displaced by large tumours, and preoperative knowledge of where these tracts are in relation to the tumour allows for safe surgical

approaches to be planned. DTI can be integrated with standard neuronavigation to allow for intraoperative estimation of the proximity of these tracts to the resection cavity. Both fMRI and DTI have been shown to change clinical decision making and surgical approach and to have a positive effect on EOR[147-149], although no randomised or head-to-head studies are available.

Awake surgery with cortical and subcortical stimulation

The most direct method for assessing function is awake craniotomy with intraoperative cortical and subcortical mapping. It is also possible, however, to do some functional monitoring with the patient asleep using electrophysiological monitoring of specific motor or sensory function. Having the patient awake in surgery allows for real-time direct assessment of brain functions, and is particularly important for speech assessment. After the craniotomy is performed, cortical mapping is undertaken in order to identify key motor, speech or sensory functions that might be present in that given brain location. The brain is stimulated with a small electrical current, and the patient asked to carry out appropriate tasks relating to that brain region. The electrical stimulation causes a temporary change or cessation of the tested function. These areas can then be safely avoided during surgery. Once deeper into the brain, subcortical stimulation is used to identify deep white matter functional tracts, often done in correlation with the DTI imaging. Once a functional area is detected, the resection is halted. A number of case series have shown the benefit of intraoperative monitoring[96,101,150-155], and its use has been particularly adopted for LGG when functional guided resections are carried out[153]. Its intraoperative use when HGG is present in movement-related brain areas was shown in a single-centre RCT to be associated with significantly better GTR rates, a lower risk of movement loss, and a better life expectancy[156]. A combined approach using DTI and intraoperative mapping is reported to increase surgical safety and increasing EOR[157-159]. Work by our group has shown that a DTI-informed approach during awake surgery reduces the occurrence and severity of postoperative problems after surgery, leading to faster recovery and shorter hospital stays[160].

1.6 Raman spectroscopy in neuro-oncology

The field of biomedical optics aims to use the various interactions of tissue with light to provide new diagnostic or therapeutic approaches. When light (electromagnetic radiation) interacts with tissue, the result can be elastic or inelastic (Raman) scattering, absorption (which can result in fluorescence if a suitable fluorophore is present) or reflection. These are shown schematically in terms of the molecular energy levels involved in Figure 1.3. These various interactions can be measured by a variety of spectroscopic approaches: elastic scattering, Raman (inelastic), absorption, fluorescence and reflectance spectroscopy, respectively. This thesis focuses on the application of Raman spectroscopy and fluorescence spectrometry.

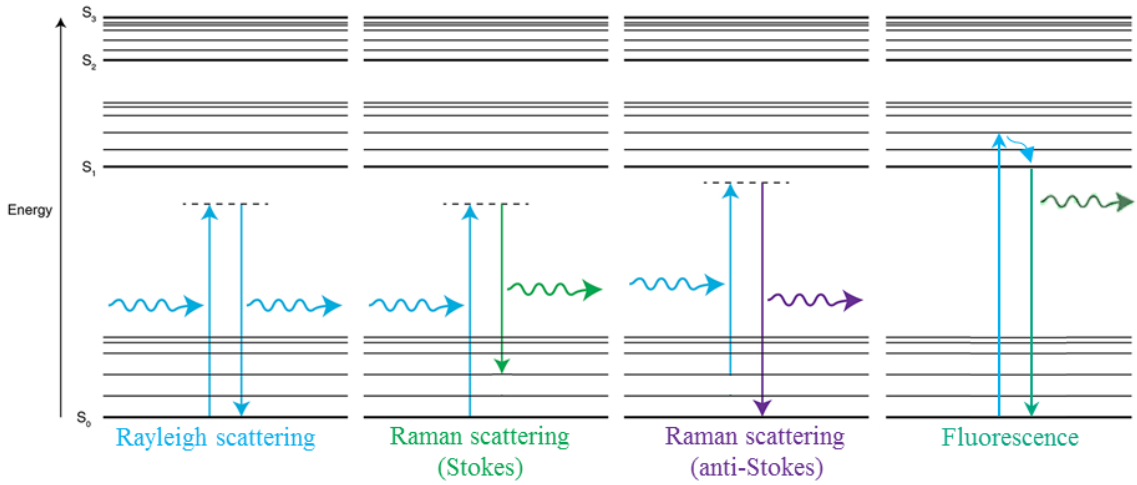


Figure 1.3 - Light interaction with a sample. Energy level diagram depicting elastic (Rayleigh) and inelastic (Raman) scatter and fluorescence. Broken lines represent virtual energy states.

1.6.1 Definition of Raman scattering

Elastic scattering, also known as Rayleigh scattering, occurs when a photon of light interacts with the sample and is scattered with its wavelength, or energy, unchanged. In a small proportion of scattering events, energy is transferred between a photon and the sample (or vice versa), yielding a scattered photon with a higher or lower energy than the incident photon (Figure 1.3). This process is known as inelastic or Raman scattering, named after C.V. Raman, who won the 1930 Nobel prize for physics for the discovery of this phenomenon[161-163]. Molecules within the tissue sample have well-defined electronic and vibrational levels. Raman scattering from a molecule is usually described as a two-step process in which the incoming photon is absorbed, exciting the molecule to a 'virtual' state, which does not

generally correspond to one of the molecular energy levels, but can be thought of as a very short-lived state of the molecule-photon system. This is followed by emission of a second photon (the Raman scattered photon) and simultaneous relaxation of the molecule to a lower-lying energy state which is different from the original molecular state. The difference in energy between the incident and scattered photons is identical to the difference in energy between the initial and final states of the molecule. The molecule can either increase its energy (Stokes scattering) or decrease its energy (anti-Stokes scattering) The work in this thesis concerns vibrational Raman spectroscopy, in which the energy levels involved are vibrational levels of the sample molecules. The photon energy shift is indicative of discrete vibrational modes (e.g. stretching, bending, scissoring) of the chemical bonds within molecules in the sample, and thus provides a qualitative measure of biochemical composition. The "vibrational fingerprint" of biological tissue recorded in a Raman spectrum reflects its chemical composition, which includes nucleic acids, proteins, lipids and carbohydrates. Stokes Raman scattering is generally more intense than anti-Stokes scattering, since most molecules are initially present in their ground vibrational state and can therefore only step up the ladder of vibrational energy levels during a Raman scattering event. Stokes scattering is therefore most commonly measured experimentally. Raman spectra are typically plotted as a function of the shift in wavenumber (wavenumber is the reciprocal of the wavelength and is directly proportional to the energy of the photon), in units of cm^{-1}, relative to the wavenumber of the excitation light. The wavenumber shift represents the loss in energy of the scattered photon. The *y*-axis is an arbitrary measure of intensity and is generally labelled in terms of "counts" or "intensity" in the present work. The Raman scattering signal is very weak compared to Rayleigh (elastic) scattering, occurring for approximately one in 10^8 incident photons.

In general, a Raman spectroscopic microscope system will consist of an excitation light source (laser), an objective to focus the excitation light onto the sample and to collect the scattered light, a wavelength filter to remove the Rayleigh scattered light, a grating to disperse the spectrum of wavelengths scattered from the sample, and a sensitive detector (most commonly, a charge-coupled device (CCD)) to give an electrical readout proportional to the intensity of the scattered light[164,165]. Raman spectrometers have undergone significant technological advances in all of the above components in order to develop them into useful analytical tools for clinical applications.

1.6.2 Advantages of Raman spectroscopy

Table 1.2 summarises a number of the attractive features offered by Raman spectroscopy which have led to significant research into its application in biomedical science. Samples require no processing, preparation or staining prior to Raman analysis, allowing for rapid and cost-effective analysis of tissue without the need for specialist staff, time-consuming laboratory protocols, or expensive consumables. Raman is also non-destructive, an important advantage for *in vivo* applications. Raman microscopes are relatively easy to use with portable, "plug and play", systems available allowing for straightforward clinical translation of developed applications. In addition, fibre optic probe-based systems also increase the potential for clinical application, particularly for intraoperative analysis of tissue.

Advantages of Raman spectroscopy
No sample preparation, processing or staining required
Non-destructive, non-invasive, non-contact
Easy to use, portable systems with clinical translation potential
Fibre optic probe-based systems available
High spatial resolution possible (< 1 μm)
Not affected by water – can be used on biological samples

Table 1.2 - Advantages of Raman spectroscopy for biomedical application

1.6.3 Applications of Raman spectroscopy in neuro-oncology and neurosurgery

There is growing interest amongst the clinical community in neuro-oncology applications of Raman spectroscopy, with two recent review articles published in clinical neurosurgery and neuro-oncology journals[166,167].

The first efforts to record Raman spectra of rat and human brain tissue were reported by Mizuno *et al.* in the early 90s[168,169]. Using frozen samples which were thawed and macerated, Mizuno *et al.* recorded

Raman spectra of normal grey matter, normal white matter, glioma, acoustic neuroma and central neurocytoma[169]. Since these first reports, research into the applications of Raman spectroscopy in neuroscience can broadly be categorised into two groups: 1) using Raman scattering as a diagnostic tool to distinguish normal from tumoural tissue and to determine different tumour types and grades; and 2) using hyperspectral Raman microscopy imaging to develop digital images of samples based on the Raman spectral characteristics. There is overlap between applications as the spectral characteristics of different brain tissue types implicated in the hyperspectral data can be used to inform diagnostic classification models and vice versa.

Studies that aim to use Raman spectroscopy as a diagnostic tool generally involve the collection of multiple Raman spectra from different tissue types, pre-processing of the raw data in order to extract the important spectral features, an assessment of the spectral characteristics which relate to the known biochemical differences between the tissue types, and then the use of machine learning algorithms to develop a classification model to classify the different tissue types into diagnostic groups. Machine learning algorithms can be supervised or unsupervised. Supervised methods, (e.g. linear discriminant analysis (LDA), support vector machines, decision tree learning) use reference knowledge, i.e. pathological class of a training set of data, to build a classification model which is then applied to a new data set (validation data set) to validate the model. Unsupervised methods (e.g. principal component analysis (PCA), cluster analysis) have no reference information and classify based on similarities and differences in the spectral characteristics alone[170].

Studies that use hyperspectral Raman imaging generally involve detailed Raman mapping of tissue samples in order to create high resolution spectral maps of tissue which are then compared with standard histological stained images. The use of false-colour maps to represent specific Raman spectral characteristics, such as peak intensities or ratios of peak intensities, allows visualisation of the chemical signatures to produce a virtual pathological image.

In both applications, studies have been undertaken using *ex vivo* animal models of glioma[171,172] and *ex vivo* human formalin-fixed paraffin-embedded (FFPE)[173] and frozen[170,174-186] normal brain and tumour tissue. *In vivo* studies have been undertaken in animal models of glioma[187] and metastases[188] and, much more recently, in humans, with intraoperative applications having been reported[189-191].

1.6.4 *Ex vivo* Raman applications

Work using *ex vivo* animal models includes that of Tanahashi *et al.*, who showed that they could distinguish infiltrative glioma from normal brain with 98.3% sensitivity and 75% specificity using a diffuse infiltrative glioma mouse model[171]. Amharref *et al.* demonstrated that Raman hyperspectral microscopy could discriminate between glioma, necrosis, oedema and healthy normal brain using frozen samples from rats inoculated with C6 glioma tumour cell line[172].

The bulk of research into Raman spectroscopy applied to brain tumours has employed *ex vivo* human tissue which has normally been cryopreserved (frozen) or FFPE. Gajjar *et al.* published the largest analysis of FFPE samples, which included 52 cases made up of normal brain, meningiomas, gliomas (high and low grade) and metastatic brain tumours. Using a PCA-LDA model they reported a statistically significant difference between the spectra of the different pathological types, as well as between the spectra of normal brain and glioma and between those of high and low grade gliomas[173]. Thawed cryopreserved samples are used by most researchers as these are thought to represent the closest approximation to fresh tissue, while avoiding the effects of formalin fixation and paraffin artefacts seen with FFPE sections[192,193]. Krafft *et al.* have published several papers on Raman analysis of cryosections, and cite variations in lipid content as the key differentiator between normal brain and glioma[170,174,175]. This group have also used hyperspectral Raman imaging of cryosections to demonstrate measurable spectroscopic and structural differences allowing for imaging of normal brain structures (meninges, white and grey matter), low grade astrocytomas, glioblastomas and meningiomas[176] as well as using Raman data to differentiate between differing degrees of glioma cell density[177]. In addition, they report the use of hyperspectral unmixing (an unsupervised learning algorithm) to give subcellular morphological and biochemical resolution with strong correlation to standard histopathological processing[178,179].

Gregory Auner's group from Detroit has focused on the application of Raman spectroscopy to the classification of paediatric brain tumours. Using fresh or frozen samples, from 29 patients, of medulloblastoma, glioma (including astrocytomas, oligodendroglioma, ependymoma), and normal brain, this group demonstrated the ability to classify the different tumour types with an accuracy of 91% - 100%[180]. High grade ependymomas could also be distinguished from low grade ependymomas, similarly

high grade from low grade glioma. A follow up study, with a similar data set and method, reported 88.9% accuracy for predicting brain tumour type[181]. A further study by this group used Raman scattering to distinguish grey matter, tumour necrosis and glioblastoma in 40 frozen tissue sections and reported a 97.8% accuracy in differentiating between these pathology types using discriminant functional analysis[182]. This group have also applied hyperspectral Raman imaging to the analysis of frozen brain sections, and have demonstrated accurate molecular imaging of normal brain and tumours, with identification of grey/white matter boundaries, glioblastoma tumour bulk, necrosis and infiltrating tumour[183,184].

Further classification models using cryosections have been reported by Zhou *et al.*[185] and Aguiar *et al.*[186]. Zhou *et al.* measured Raman spectra of frozen samples from seven patients with a range of pathologies including high grade (WHO III) meningioma, low grade (WHO I) meningioma, normal brain meninges, glioblastoma, acoustic neuroma and pituitary adenoma. A diagnostic model using PCA and support vector machine was built to classify the meningeal pathologies into normal meninges, high grade meningioma and low grade meningioma with a sensitivity of 90.9% and specificity of 100%[185]. Aguiar *et al.* used single frozen samples of normal cerebellum, normal meninges, glioblastoma, medulloblastoma and two samples of meningiomas to build a PCA classification model which demonstrated 80.0% - 89.7% sensitivity and 97% - 99% specificity for predicting the correct tumour class[186].

With the exception of the hyperspectral imaging studies that have a clear aim of providing a virtual pathological image based on Raman spectral data, most of the classification studies reported above have classified a rather eclectic group of tumour types, which would appear to have been dictated by the availability of tumour types in the source brain bank rather than guided by a clear clinical or neuropathological question. The low number of samples of each pathology type is also of concern, resulting in an over estimate of classification model performance as the effect of intra-pathological variation is not accounted for when only one or two samples of each pathology group are used. These limitations reflect the early, exploratory and proof-of-concept stages of this research field in its present state rather than implying a severe criticism of the studies performed to date. Research in this area has started to yield a spectral library of different brain tumour types, characterising them in terms of their

Raman "fingerprint". However, to make serious progress towards clinical application of Raman spectroscopy, it is important to consider the specific clinical or pathological applications when designing future studies. It is also essential to target applications that play to the advantages of the technique (i.e. non-destructive, no sample preparation, rapid technique). When considering the use of Raman spectroscopy as an *ex vivo* diagnostic tool, it is important to know which tumour subtypes a surgeon, oncologist or neuropathologist would need to rapidly differentiate in order to influence the surgical or oncological treatment strategy. That is, what diagnostic information can Raman spectroscopy offer that might influence patient management and cannot wait for more complex, lengthy, diagnostic procedures? More recent *ex vivo* studies have started to focus on these important clinical or neuropathological questions. With the 2016 revision of the WHO classification of CNS tumours and the incorporation of molecular genetics, the need to build brain tumour Raman classification models using these new diagnostic criteria is essential. Uckermann *et al.* is, to date, the only paper to have approached this issue, when reporting their efforts to differentiate IDH-mutant from IDH-wildtype gliomas[194]. In their study, 36 human glioma samples were obtained during surgery, snap frozen, sectioned, and subjected to Raman analysis. The detailed pathological diagnosis of each tumour was not reported, but samples consisted of two oligodendrogliomas, two glioblastomas, and several different astrocytomas, with a total of 12 IDH-wildtype and 24 IDH-mutant tumours. A number of differences were identified between the Raman spectra of the IDH-wildtype and IDH-mutant gliomas, which mainly corresponded to differences in DNA, protein and lipid peaks. Selected Raman peaks were used to test a classification approach, based on Mahalanobis distances, for prediction of IDH-mutant status, giving a correct classification rate of 89% on independent validation. The promising results of this study suggest that Raman spectroscopy can be used to classify gliomas in terms of their genetic subtypes. This application is explored further in the work presented in this thesis.

1.6.5 *In vivo* Raman applications

In vivo Raman analysis of tumour and normal brain was first reported in a mouse model by Kirsch *et al.*[188]. Mice were inoculated with either a human or murine melanoma cell line, and Raman spectral maps of the cortical surface were recorded using a fibre optic probe via a cranial window made in the anaesthetised mouse. Using cluster analysis, it was possible to distinguish tumour from normal

surrounding brain for both cortical and subcortical tumours. Beljebbar *et al.* reported the *in vivo* detection of glioma, using C6 glioma inoculated rats, and showed similar abilities to differentiate between tumour and non-tumour[187]. Both of these studies are limited in their spatial resolution due to the fibre optic setup used but provided proof-of-concept for the use of *in vivo* Raman spectroscopy to distinguish tumour tissue from normal tissue.

Fredrick Leblond's group in Montreal were the first to undertake *in vivo* Raman spectroscopy of human brain tissue. A preliminary study reported the development of a hand-held Raman probe used in 10 patients to differentiate necrotic tissue from vital tissues (which they define as including tumour and normal brain) with an accuracy of 87%, sensitivity of 84% and specificity of 89%[189]. The group then report the use of their Raman probe system for intraoperative assessment of tissue as tumour versus non-tumour[190]. The case series consisted of 17 patients with a mix of tumour pathologies (three WHO II astrocytomas, one WHO II oligodendroglioma, one WHO III astrocytoma, one WHO III oligodendroglioma, one oligoastrocytoma[v], eight glioblastomas and two metastases). Brain samples were also taken from areas around the tumour bulk and designated by the neuropathologists as either normal brain or infiltrating tumour. Using a boosted trees machine learning algorithm, they were able to distinguish normal brain from both bulk tumour and infiltrating tumour with an accuracy of 92%, sensitivity of 93% and specificity of 91%. In an attempt to estimate the threshold of cancer cell density that the Raman probe correctly classified as cancer, a subset of samples, deemed by the neuropathologists to have the lowest tumour cell burden, were analysed and it was found that all false-negative Raman spectroscopy classifications corresponded to samples with <15% cancer cell burden. A critique of this study is given in Chapter 4, but nonetheless this first-in-man case series shows the significant potential of Raman spectroscopy to become a surgical adjunct to aid resection of brain tumours. In a follow-up study using the same probe setup and Raman classification model methodology, Leblond's group compare the detection of tumour by Raman probe with the tumour identified on preoperative MRI scan. Using neuronavigation to correlate the location of the Raman measurements on the MRI scans, they

[v] Oligoastrocytoma is an out of date diagnosis which was given to tumours with mixed astroglial and oligodendroglial features, based on histopathological evaluation. This has now been superseded by molecular testing, as described earlier in this chapter, which confirms astro- or oligodendroglial origin.

demonstrate that Raman spectroscopy detects invasive cancer up to 3.7 cm and 2.4 cm beyond the T1-contrast-enhanced and the T2/FLAIR MRI boundary respectively[195].

There has been some *in vivo* work on incorporating Raman analysis into a neurosurgical biopsy probe, enabling the use of Raman spectroscopy to differentiate between tumour, necrosis and normal brain, with the suggestion that this would increase the diagnostic yield of biopsy samples[191]. The problem of diagnostic yield is not significant enough to warrant the uptake of this technology into routine neurosurgical practice, and therefore this should not be the focus of further research unless other applications can be found.

In general, the limited work carried out to date on human *in vivo* applications of Raman spectroscopy has correctly focused on interrogating the tumour boundary and building models of tumour versus non-tumour samples. The first *in vivo* reports are very promising, but much more evidence, validation and testing is required before this technology will be widely used by the neurosurgical community.

1.6.6 Coherent Raman spectroscopy

So far in this chapter, the type of Raman spectroscopy discussed simply measures the scattered photons that have spontaneously undergone Raman scattering, known as spontaneous Raman spectroscopy. The main limitation of spontaneous Raman spectroscopy is the small number of photons that undergo inelastic scattering and thus the weak Raman signal. Advances in Raman spectrometer instrumentation have certainly improved the signal to noise ratio achievable for spontaneous Raman spectroscopy[165], but significant enhancement of the Raman signal has been achieved with the development of coherent Raman spectroscopy (CRS). CRS uses a second excitation, or pump, light source to coherently drive the vibrational frequency of Raman-active modes, thus increasing the Raman signal from these modes. Two major forms of CRS have been developed and used in biomedical sciences – coherent anti-Stokes Raman scattering (CARS) and stimulated Raman scattering (SRS). Both techniques can increase the Raman signal by a factor of approximately 10^4. This has the advantage of significantly increasing the Raman peak resolution allowing for better differentiation of tissue types both when using diagnostic models and virtual histopathology. CARS has been successfully used to image normal brain, glioma, and metastases in *ex vivo* animal[196-199] and human samples[199]. SRS has several advantages over CARS and has been developed as a biomedical imaging modality with the capability of providing highly detailed virtual

histopathological images of normal brain and tumour[200,201]. SRS has been applied *in vivo* in glioblastoma rat models, demonstrating that SRS microscopy could identify tumour infiltration in areas that appeared normal by eye, suggesting a possible surgical application[202]. Furthermore, SRS microscopy was carried out *ex vivo* on fresh, unprocessed surgical samples from 22 patients, and demonstrated the ability to detect glioma tumour infiltration with 87.3% sensitivity and 87.5% specificity on independent cross validation[203].

Despite these promising reports of CRS, and particularly SRS, applications in neuro-oncology, there are significant barriers to translation into the clinical setting. The biggest hurdle is related to the instrumentation needed. SRS requires two laser pulses that are temporally overlapped by less than the pulse duration (<100 fs) and spatially overlapped by less than the focal spot of the lasers (<100 nm). Achieving these conditions in a system to be used in a clinical setting is a significant challenge. Orringer *et al.* report the first use of a fibre-laser-based SRS microscope, which they engineered for use in the clinical setting[204]. This SRS microscope was used to image *ex vivo* fresh tissue samples from 101 neurosurgical patients giving impressive virtual histological images – "stimulated Raman histology" – with remarkable concordance with conventional neuropathological staining. With regards to diagnostic classification, they reported a 90% accuracy in predicting the correct tumour diagnosis, although this model was only built using data from 30 patients. This group later reported a similar application of their SRS microscope for the *ex vivo* analysis of fresh paediatric brain tumour samples[205]. Once again there was near perfect concordance between the pathological diagnosis achieved using the stimulated Raman histological image and that achieved with the standard H&E image. Using machine-learning diagnostic classification, they report 93.8 % accuracy for predicting normal versus tumour tissue type and 89.4% for predicting low-grade versus high-grade tumour. They did not attempt to classify in terms of tumour pathological type.

The evidence would suggest that SRS is a superior technique for virtual histology of brain and tumour compared to hyperspectral spontaneous Raman spectroscopy, and currently performs on a par with spontaneous Raman spectroscopy as a diagnostic tool for pathological classification. Spontaneous Raman methods have the advantage of being much simpler in terms of engineering, which makes them more reliable and more easily translated into the clinical setting as well as being lower in cost. Currently

SRS requires a microscope objective setup, and is not compatible with fibre optic-based systems. This also gives spontaneous Raman approaches the significant advantage of offering possibilities for intraoperative use with a hand-held probe, as described earlier.

1.6.7 Fourier-transform infrared spectroscopy

Fourier Transform Infrared Spectroscopy (FTIR), like Raman spectroscopy, is a form of vibrational spectroscopy. Unlike Raman scattering, which involves absorption and re-emission of a photon, IR spectroscopy involves a simple absorption measurement in which the detector measures the absorbance of infrared radiation by a sample. Each sample absorbs different amounts of each frequency of IR radiation, resulting in a "chemical fingerprint" in the form of the infrared spectrum. The application of FTIR spectroscopy *in vivo* has been challenging historically due to the strong absorption of mid-infrared radiation by water, which results in masking of the biochemical spectra by large water peaks and thus limits the use of FTIR to dried samples. This problem has largely been overcome through employing attenuated total reflectance (ATR) probes. ATR-FTIR spectroscopy requires little or no sample preparation and is largely unaffected by strong infrared absorbers such as water. There are many comparisons between biomedical applications of Raman and ATR-FTIR, and the two approaches should generally be thought of as complementary rather than competing technologies[206-209]. In general, FTIR measurements yield spectra with higher signal to noise ratios and faster sample mapping compared with Raman spectroscopy, but have inferior spatial and spectral resolution[207]. FTIR spectroscopy has been shown to accurately differentiate normal brain from glioblastoma frozen samples, as well as to differentiate between varying grades of glioma, with accuracies of between 93.6% and 100%[210]. Similarly high differentiation between glioma subtypes was found using FFPE samples[173]. A study using frozen sections of brain metastases (from renal, lung, colorectal and breast primaries) demonstrated the ability of FTIR to accurately classify the metastasis according to the primary cancer origin and to differentiate metastasis from normal grey and white matter[211]. Uckermann *et al.*, similar to their publication described above on the application of Raman spectroscopy for the differentiation of IDH mutation status of gliomas, report using ATR-FTIR on fresh glioma samples to give correct IDH mutation status classification in 88% of the tumour samples[212]. In general, the diagnostic classification performance of FTIR spectroscopy in these *ex vivo* studies is comparable to similar reports using Raman

spectroscopy. With regard to *in vivo* use, the development and application of fibre-optic probe-based FTIR systems has been slower than for Raman probe systems, although there has been some work using FTIR probes for *in vivo* detection of skin, bladder and colorectal cancer[213]. One disadvantage of ATR-FTIR probes is that the ATR crystal needs to come into contact with the tissue with a small amount of pressure, which may not be possible or desirable during neurosurgical application. The fibre probes are also sensitive to ultraviolet light and humidity, limiting their potential for *in situ* application[213]. The penetration depth of ATR probes, typically just 2 or 3 μm, is also a limiting factor in the surgical application of FTIR methods, as is the current size of commercially available ATR probes, which renders them inapplicable for endoscopic applications[214]. Despite these limitations, there is no doubt that the design and instrumentation of ATR-FTIR probe-based systems will improve to match that of Raman devices, and thus development of these two vibrational spectroscopy techniques should be continued in synergy.

1.7 Fluorescence in neuro-oncology

Fluorescence occurs when a photon of light is absorbed by a sample and electrons are excited to electronic energy levels. After rapid vibrational relaxation within the electronically excited state, the molecule returns to the electronic ground state by emitting a photon at a longer wavelength (lower energy) than the incident light (Figure 1.3). Within the field of neuro-oncology surgery, the only routinely used fluorescence tool is 5ALA-induced fluorescence, in which PpIX fluorescence is used to guide resection of HGG. The main clinical evidence for and application of this approach have already been covered in Section 1.5.2.

1.7.1 Improving the sensitivity of 5ALA-induced fluorescence

There has been a body of work investigating the ability of PpIX fluorescence measurements to detect tumour, together with attempts to increase the detection sensitivity for fluorescent tumour[215]. In typical fluorescence-guided surgery, the standard wide-field operative microscope is used to visualise the fluorescence by switching to a blue light source and visualising the pink fluorescence via the microscope eyepiece. Several studies have looked at the accuracy of using visual fluorescence, which is often given a visual intensity score related to the degree of fluorescence, to predict the presence of

tumour[131,132,134,216,217]. These studies generally demonstrate that intense fluorescence is a strong predictor of the presence of tumour, but that the absence of fluorescence is a poor indicator for the lack of tumour. One typical study reported that the strongest visible fluorescence correlated with the presence of tumour 97.4% (positive predictive value) of the time, and that absence of visual fluorescence correlated with the absence of tumour only 37.7% (negative predictive value) of the time[217]. Attempts to improve the detection of fluorescence below the visual threshold have included the application of fluorescence probes[134,135,218] and fluorescence microscopes[219], with quantitative techniques allowing for concentrations of PpIX to be measured from the acquired spectra[135,219-221]. Intraoperative fluorescence probes can detect fluorescence in HGG well beyond the visual threshold, improving the negative predictive value up to 77%[135]. These fluorescence probe systems have also shown the possibility of detecting the low levels of fluorescence in LGG, which generally do not fluoresce enough to reach the visual threshold[222].

Despite the increased sensitivity for the detection of fluorescence using fluorescence probes or microscopes, their significant disadvantage is the interruption these techniques cause to the surgical flow. The huge advantage of the wide-field microscope systems currently in use is that the surgeon can switch between white light and fluorescence without the need to stop operating. Surgical resection continues, with the fluorescence light giving real-time feedback as to the resection of fluorescent tumour. Repeat interrogations of a small area of the surgical field with a probe-based system prolongs and interrupts the surgery, and thus application of such a device needs to be carefully considered to make sure the benefit outweighs this disadvantage.

Our group has developed the use of endoscopic resection of gliomas, with the benefits of a key-hole surgery approach, better visualisation of deep tumours, and fast postoperative recovery[160,223,224]. We also have the largest series of endoscopic resections of gliomas using 5ALA-induced fluorescence, with only a few case series previously published in the literature[225,226]. In our series of over 40 patients, we report visualisation of fluorescence with the endoscope that was not seen with the standard surgical microscope (manuscript under review). The use of an endoscope may be a relatively simple way of increase the visual detection of fluorescence using currently available surgical equipment.

1.8 Project background

This thesis reports on a project to investigate applications of Raman spectroscopy and fluorescence spectrometry in glioma surgery and neuropathology.

A new collaboration was formed as part of this project between the University of Oxford Departments of Chemistry, Neuropathology, and Neurosurgery, and an industry collaborator, Renishaw Plc.(Wotton-under-Edge, UK). The intention was to establish clinical spectroscopy research within the Departments of Neuropathology and Neurosurgery.

Renishaw are industry experts in Raman spectroscopy applications, and the focus of the collaboration was provision via a loan of a bench-top Raman spectrometer to the Departments of Neuropathology and Neurosurgery with a view to analysing *ex vivo* fresh brain samples, straight from neurosurgery theatres. An intraoperative fibre-optic probe-based Raman system may well be a long term aim for this collaboration, but was not within the remit of this project, as proof-of-concept using Renishaw's current instrumentation was first required. Prior to using Raman spectroscopy on fresh *ex vivo* tissue, a significant amount of protocol development was carried out.

With regards to fluorescence spectrometry, our group are developing new methods of enhancing visualisation of 5ALA-induced fluorescence which require testing in a laboratory setting prior to trialling in the operating theatre. To this end, this project involved the development and calibration of a new laboratory-based fluorescence measurement setup and characterisation of a stable fluorescence standard to allow testing of new fluorescence visualisation devices. As mentioned in Section 1.7.1, our neurosurgical group has developed the use of endoscopic 5ALA-induced fluorescence-guided surgery for gliomas. The laboratory-based fluorescence setup was used to test the hypothesis that greater visual fluorescence is detected using the endoscope compared with the microscope.

The neuro-oncology applications of Raman spectroscopy and fluorescence spectrometry sit very well together. Intraoperative fluorescence probes have been shown to improve fluorescence detection and therefore developing a fluorescence probe-based system, alongside the Raman application development, was important to allow for future comparison of these techniques. Moreover, the current gold-standard for the resection of HGG is to use 5ALA-induced fluorescence-guidance. It is therefore important that

any new technique developed to detect intraoperative tumour margins, such as Raman spectroscopy, must be compared with this gold standard.

Future work will build on the foundations established by this current project, and the development of both Raman and fluorescence applications will allow for dual testing of these modalities when future projects move on to intraoperative evaluation of these or other devices.

1.9 Project aims

The aims of this project were:

1. to apply Raman spectroscopy to intraoperative and neuropathology analysis of gliomas both in terms of genetic subclassification and tumour/normal boundary;
2. to develop a fluorescence spectrometry laboratory setup to allow for testing of enhanced fluorescence visualisation methods and to enable future comparison of fluorescence with Raman spectroscopy for the detection of tumour transition zone.

1.10 Thesis outline

Chapter 2 describes Raman spectroscopy protocol development, involving optimisation of both the sample preparation methods and the Raman measurements. In Chapter 3, several variables which may affect the Raman spectra are investigated. This work includes an evaluation of the spectral changes induced by exposure of the samples to the Raman laser beam. Methods developed in order to achieve a compromise between good quality spectra and limited spectral change due to laser exposure are reported. The main clinical application of Raman spectroscopy investigated in this project – the intraoperative genetic classification of gliomas– is reported in Chapter 4. Chapter 5 details the development of a fluorescence standard and laboratory-based fluorescence setup, which is used to evaluate the fluorescence detection thresholds of the operative microscope and endoscope. Chapter 6 brings together Raman spectroscopy and visual fluorescence detection by reporting on a preliminary study that evaluates the relative ability of Raman spectroscopy and visual fluorescence to differentiate infiltrating tumour from normal brain.

Chapter 2. Optimisation of sample preparation and Raman spectrometer configuration

2.1 Introduction

Optimisation and standardisation of protocols is essential prior to undertaking experiments that aim to answer specific research questions. There have been many attempts by the vibrational spectroscopy community to form a consensus on the standardisation of sample preparation and spectrometer setup used for clinical applications of Raman spectroscopy[164,209,227,228]; however there is still debate surrounding this issue. Changes in sample preparation or spectrometer configuration can have a significant effect on the Raman spectra, and it is therefore of key importance that these do not change throughout the data collection. Once a protocol has been developed that produces good quality spectra it is essential that this is adhered to throughout. This is particularly important when machine learning predictive models are to be employed, as spectral changes due to variation in protocol may be greater than the changes due to pathology type, thus having detrimental effects on model performance. Protocols developed must be easily reproducible in order to maintain quality control and consistency. They must also be integrated into the current clinical diagnostic workflow in order to aid future translation into the clinical setting[209].

The aim of this preliminary methodological study was to establish the optimum sample preparation and Raman spectrometer configuration needed to acquire high quality Raman spectra of different brain pathology types. With clinical translation in mind, tissue preparation methods were chosen to reflect current protocols used in the Department of Neuropathology. Pathology types tested included tumour subtypes that would be used in future studies. The objective was to confirm that good quality spectra could be obtained from samples prepared using the existing neuropathology protocols and tissue preparation types that were available within the Oxford Brain Bank.

With regards to the spectrometer configuration, the main variable investigated was the excitation laser wavelength. Preliminary data from the *Rapid molecular genetic classification of gliomas using Raman spectroscopy* study (Chapter 4) were used for this. All Raman spectrometer equipment was provided by the industry collaborator, Renishaw Plc. The results of this preliminary study were used to inform the

decision as to the specification for a loan spectrometer instrument that was later installed onsite in the Department of Neuropathology.

2.2 Method

2.2.1 Sample preparation variables

The follow sample preparation variables were tested:

a) pathology type:
    1. normal brain (including normal cerebral and cerebellar cortex),
    2. WHO grade IV astrocytomas with IDH mutation (IDH-MUT),
    3. WHO grade IV astrocytomas without IDH mutation, IDH-wildtype (IDH-WT),
    4. tumour necrosis;
b) tissue preparation:
    1. cryosection (Cryo),
    2. formalin-fixed paraffin-embedded sections (FFPE);
c) substrate used for mounting sample:
    1. calcium fluoride (CaF_2),
    2. stainless steel (SS),
    3. fused silica (SiO_2).

Suitable tissue samples were identified from the Oxford Brain Bank by the Consultant Neuropathologist. The normal brain sample was of cerebellar cortex from a normal brain post-mortem donation. The tumour samples were both WHO grade 4 astrocytomas (glioblastomas), one with a proven IDH mutation (IDH-MUT) and one without an IDH mutation, IDH-wildtype (IDH-WT). Details of immunohistochemistry and targeted genetic sequencing used to determine IDH status of these tumours can be found in Chapter 4. The necrosis was from the IDH-WT glioblastoma sample which contained a large region of tumour necrosis. For all pathology types both cryopreserved and FFPE tissue was available.

Cryopreserved samples were thawed from -80 °C to -40 °C and then 10 µm thick cryosections were cut using a cryotome. Parallel cut sections were mounted on CaF_2, SS, SiO_2 and glass slides. The glass slide

samples underwent standard haematoxylin and eosin (H&E) staining and were used by the Neuropathologist to identify areas of interest (i.e. normal brain, tumour, necrosis) on the slide. The remaining slides destined for Raman analysis were kept in a slide box and returned to the -80 °C freezer. Samples were packed on dry ice for transportation to Renishaw laboratories, and were allowed to thaw for 15 minutes prior to undertaking Raman measurements.

FFPE samples were sectioned at 6 µm thickness using a microtome and mounted on CaF_2, SS, SiO_2 and glass slides. The slides were baked in a 50 °C oven to encourage section adherence to the slide. Some preliminary testing of the optimum baking time was performed to balance the need for sample adherence with any possible effects on the recorded spectra. Initially, baking time was determined by inspecting the slides for residual water at 10 minute intervals, with 30 minutes of baking found to be enough to remove most of water from the slide. Others have reported no change in Raman spectra after 24 and 48 hours of baking at 50 °C[229] thus it was assumed that 30 minutes of baking would be unlikely to have any effect on the Raman spectra. The glass slide sample underwent standard protocol H+E staining and was used by the neuropathologist to identify areas of interest on the slide. The remaining slides were dewaxed using the Department of Neuropathology standard protocol: 3 x 5-minute baths of Histoclear (National diagnostics, USA) followed by 3 x 10 dips in 100% isopropyl alcohol (IPA). They were then rehydrated (10 dips in 90% IPA followed by 10 dips 75% IPA, 10 dips 50% IPA and 10 dips in distilled water) and allowed to air dry before being stored in a slide box until the time came for Raman analysis. It is recognised that the dewaxing process does not completely remove all the wax from the tissue and can result in wax artefacts in the Raman spectra. The efficacy of Histoclear and other dewaxing agents were evaluated by O Faolain *et al.* who demonstrated that residual wax gave a significant Raman peak contribution, namely at 1062 cm^{-1}, 1296 cm^{-1} and 1441 cm^{-1}[1230]. O Faolain *et al.* reported that wax artefacts persisted even after immersion in Histoclear for 18 hours. Similar results were seeing with other dewaxing agents including xylene and hexane as well as head-mediated antigen retrieval methods of dewaxing. Histoclear was the dewaxing agent of choice in this project because of its routine use in our, and many other, Departments of Neuropathology.

Raman grade CaF_2 and SiO_2 slides (Crystran Ltd, UK) were used in the study. Mirrored SS slides were made in house by Renishaw Plc. The glass slides used (SuperFrostPlus™, Thermo Scientific™, UK)

were of the type routinely used in the department. These glass slides come pre-coated, giving them a positive charge in order to encouraging tissue, generally negatively charged, adherence to the slide. Poor tissue adherence to the CaF_2, SiO_2 and SS slides was a problem in this study and therefore a slide coating protocol was developed. This was based on a standard coating protocol used for glass slides using APTES ((3-Aminopropyl)triethoxysilane, Sigma-Aldrich, UK). Details of the coating protocol can be found in Appendix A.

2.2.2 Raman spectrometer configuration

The key spectrometer variable that was tested was the excitation laser used. The two available laser sources used with the spectrometer had the following properties:

1. 785 nm wavelength, maximum power 195 mW,
2. 830 nm wavelength, maximum power 170 mW.

Any difference observed in the signals generated using the two lasers could originate either from the different excitation wavelengths or the slightly different laser powers, or some combination of the two. The laser powers were not matched for comparison of wavelength alone because it was felt that future experiments would likely use full laser power regardless of the laser chosen.

2.2.3 Raman spectroscopy measurements

Raman measurements were performed with an inVia™ Raman spectrometer (Renishaw Plc, UK). Data acquisition parameters were controlled using WiRE v4.3 software (Renishaw, UK). Two methods of acquiring Raman measurements were used.

1. **Single point spectra**: the excitation laser is exposed onto one single point on the sample and the Raman spectra recorded from this one location.
2. **Area mapping**: several spectra are collected from a defined area on a sample. There is a change in the x-axis and/or y-axis coordinates after every exposure. Within the acquisition software, there are multiple settings for the number of spectra and pattern of sample scanning when using an area map. The simplest is a **point-map** where a grid of spectra are collected within a rectangular area. The distance between each measurement is defined by the x-axis and y-axis step size (μm). Figure 2.1 shows part of the WiRE software point map setup window with an x-

axis step size of 100 µm and z-axis step size of 50 µm giving 66 spectra distributed in a 6 by 11 grid within the 0.5 mm^2 area. Other, more complex mapping options are available, such as StreamLine™ mapping which is described in Chapter 3.

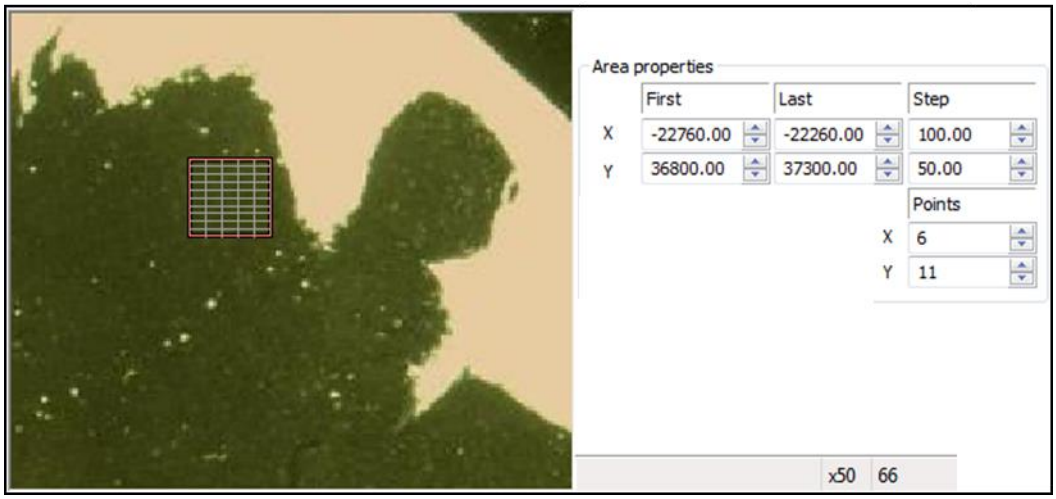


Figure 2.1 - Cryosection and FFPE point map setup. Point map setup window in WiRE software. X and Y coordinates, step size (µm) and the resulting number of points in X and Y (and total, 66) are shown. Image shows sample at x50 magnification and location of point map. (Other menu options have been masked out for clarity)

In general, the best quality spectra are achieved by increasing exposure time. Exposure time is defined as the length of time the excitation laser is exposed onto a given point on a sample and the Raman signal is collected. The exposure time is limited by saturation of the CCD detector's output. This occurs when the amount of charge generated by the Raman scattered photons incident on the detector reaches the limit of the detector's digital output. Increasing the exposure time beyond this point will produce a distorted spectral output that is flagged-up by the software. Exposure time can be split into several accumulations, for example a 2 s exposure time could be made up of four 0.5 s accumulations or two 1 s accumulations. The advantage of this is that it allows greater overall exposure with less chance of CCD saturation, although this is at the cost of increased CCD readout noise[vi]. On the inVia™ system CCD saturation occurred at around 60, 000 counts and therefore the exposure times were set to give a maximum count of approximately 50, 000.

For the comparison of different pathology types, tissue preparations and substrates, single point spectra were compared. Multiple single spectra were collected from different locations on the section. Between

[vi] Readout noise is the noise that occurs on every readout of the CCD as a result of the conversion of the electrons to a voltage. Multiple accumulations will give multiple conversions and therefore cumulative readout noise.

each measurement the white-light microscope image (for examples of low magnification white-light image see Figure 2.1 and Figure 2.2b) was viewed to confirm that the sample was in focus and that there was no thermal damage. The optimum exposure time was chosen to maximise signal to noise ratio, used as a measure of spectral quality, whilst avoiding CCD saturation or tissue thermal damage. When comparing variables, the same exposure times were used in each case to allow comparison.

For the comparison of laser type used, the first 30 FFPE samples collected in the *Rapid molecular genetic classification of gliomas using Raman spectroscopy* study were measured using both 785 nm and 830 nm excitation lasers. These data were used to evaluate the differences between spectra recorded with the two excitation lasers. This preliminary work was also carried out on the inVia™ Raman microscope and the exposure time was 4 s.

2.2.4 Data analysis

Renishaw WiRE 4.3 software was used for data analysis. Many of the spectra reported here are raw spectra, plotted without preprocessing in order to show the quality of the raw spectra and the effects of different variables that might otherwise be masked by preprocessing steps. Preprocessing used in the *Rapid molecular genetic classification of gliomas using Raman spectroscopy* study is described in detail in Chapter 4 and is not relevant here for the comparison of laser wavelengths. For the point spectra the following preprocessing steps were used in some cases:

1. Spectra were truncated to include only the region between 400 and 1850 cm^{-1} wavenumbers. The range between 1560 to 1578 cm^{-1} was also excluded, since it is affected by the Raman band of oxygen.
2. Cosmic rays[vii] and CCD saturated spectra were identified visually by the user and that spectrum excluded.
3. Baseline correction was carried out using the WiRE "Subtract baseline" function using a polymonial of order 3.

[vii] Cosmic rays are high energy particles from outer space that produce a shower of particles when they hit the atmosphere. These particles generate artefacts when they hit the CCD detector which are typically narrow, high intensity, Raman bands.

4. Normalisation was carried out using the WiRE "normalisation" tool whereby the intensity values were modified so that they vary between 0 and 1.

The assessment of spectral quality was performed qualitatively by subjective review of spectra and resolved Raman peaks. Quantitative assessment was used to compare spectral quality between different variables. Signal to noise ratio (SNR) of the spectra was used as a measure of spectral quality and area under the spectrum (integration) as a measure of amount of background intensity. These were calculated using the software's inbuilt SNR calculator (noise region: 930 to 980 cm^{-1}, signal region: 980 to 1020 cm^{-1}) and integration tools (integration region: 800 to 1700 cm^{-1}). Using this SNR calculation method a SNR of greater than 10 was considered acceptable and anything under 5 considered poor quality.

2.3 Results

2.3.1 Effect of pathology type on spectral characteristics

To investigate the dependence of the Raman spectra on pathology type, tissue spectra were compared for cryosectioned samples prepared on stainless steel substrates. Raman peaks were easily resolved in the raw spectra on all pathology types using both 785 nm and 830 nm laser wavelengths. Figure 2.2 illustrates an IDH-WT glioblastoma FFPE sample with regions of normal brain, tumour and necrosis. The parallel H&E section and the unstained section on SS used for Raman measurements are shown, along with typical raw spectra acquired with 785 nm and 830 nm excitation wavelengths for regions identified as normal, tumour, and necrosis. The prominent Raman peaks are also labelled and correspond to known Raman peaks of brain tissue reported in the literature[176,182,231-233]. Biomolecule assignment to these peaks can be seen in Table 4.3 in Chapter 4. In these raw spectra, more peaks are resolved with the 785 nm laser compared to the 830 nm laser. This may be due to the higher power of the 785 nm laser but even when the power of the 785 nm laser was reduced to match the 830 nm laser, there was still better peak resolution with the 785 nm laser. It is therefore likely that this is due to the fundamental principle that Raman signal intensity is inversely proportional to the laser excitation wavelength[234,235]. Similar quality raw spectra were recorded of different pathology types using cryosectioned tissue.

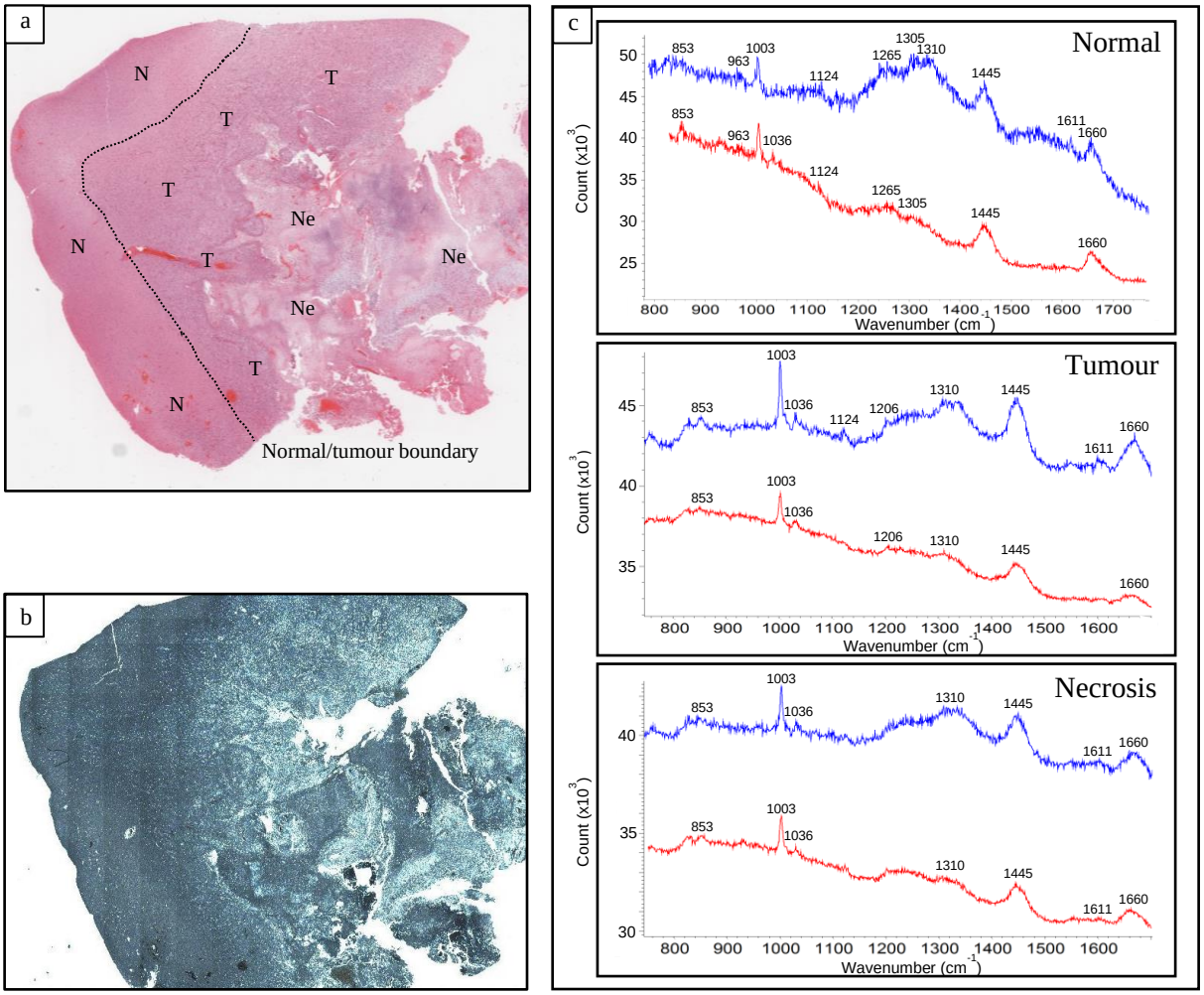


Figure 2.2 – Raw spectra from FFPE tissue sample. a) H&E stain and b) unstained sections of IDH-WT tumour with regions of normal brain (N), tumour (T) and necrosis (Ne); c) typical raw spectra for each tissue type using 785 nm (blue) and 830 nm (red) excitation lasers. Spectra have been off set for clarity.

2.3.2 Effect of tissue preparation on spectral characteristics

To investigate the effect on the Raman spectra of the tissue preparation both cryosections and FFPE sections of normal cerebellum were analysed. Cryosections have the advantage of being completely unprocessed as they are essentially frozen-fresh tissue compared to FFPE sections which have the structural changes associated with formalin fixation (cross linking of proteins) and paraffin embedding (loss of lipid). Practically, cryosections require -80 ºC storage and transportation on dry ice prior to Raman measurements, and only have a short lifespan once thawed. FFPE sections posed no such challenges, making them much more straightforward to use. Despite dewaxing the FFPE sections, a wax signal was observed in some of the FFPE samples mounted on CaF_2 slides for all pathology types (Figure 2.3).

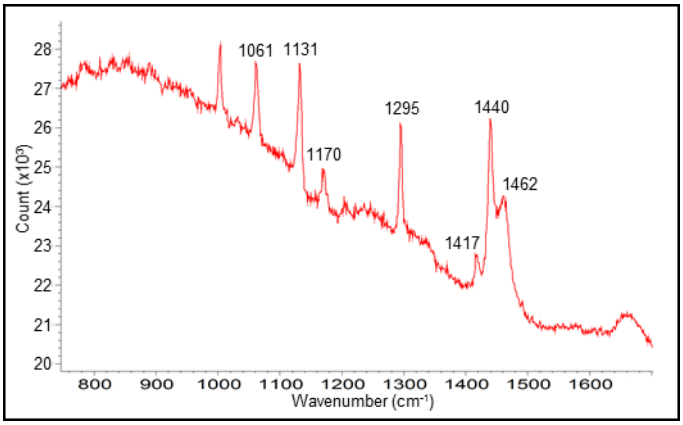


Figure 2.3 – IDH-MUT tumour on CaF_2 slide, labelled peaks represent wax signal (830 nm laser).

There were no wax peaks seen in the raw spectra when stainless steel or fused silica slides were used. Using SS, peaks were easily resolved using both cryosections and FFPE sections. As an example, Figure 2.4 illustrates the similarity in spectra and SNR for a cryosection and FFPE section of normal cerebellum. The white light microscope images in Figure 2.4 also illustrate the superior histological image quality achieved with FFPE compared to cryosection images. This makes correlation with parallel H&E sections much easier with FFPE sections.

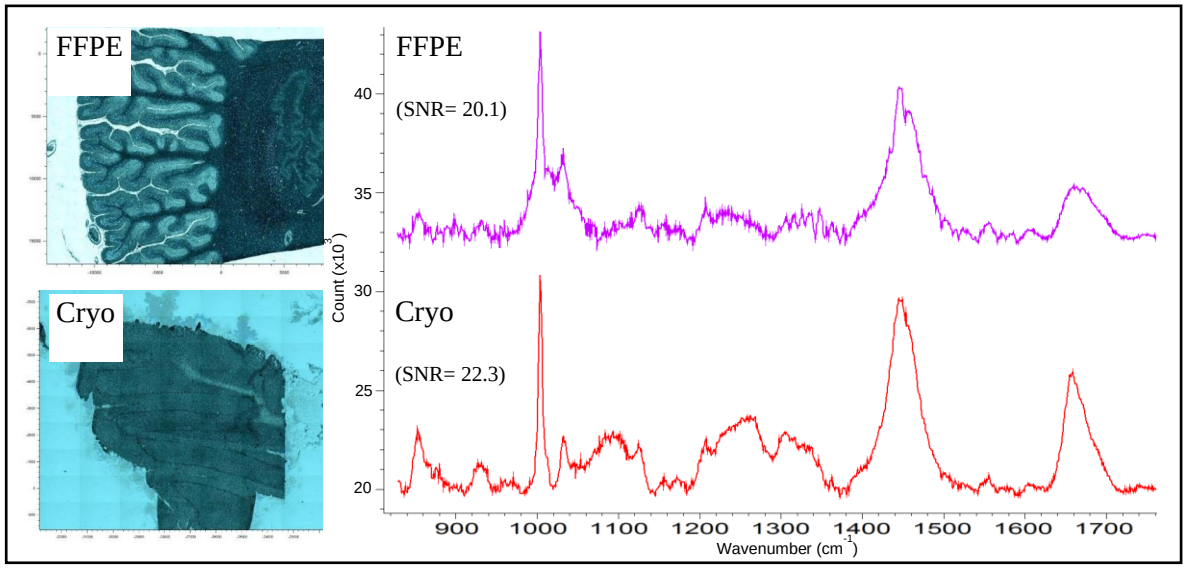


Figure 2.4 – Cryosection versus FFPE section. White light microscope images and characteristic spectra (baseline corrected (polynomial 3)) of normal cerebellum FFPE section and cryosection mounted on SS substrate using 830 nm laser and 5 s exposure time. SNR shown in parenthesis. Spectra are offset for clarity.

2.3.3 Effect of substrate on spectral characteristics

As noted previously, samples were mounted on either SS, CaF_2, or SiO_2 slides for Raman analysis. Each mounting substrate made a characteristic contribution to the background signal; this was substantially greater for CaF_2 and SiO_2 and minimal for SS (Figure 2.5). The SiO_2 background affected a large

wavenumber range, and SiO_2 was therefore deemed unsuitable as a substrate and was not investigated further. The area under spectrum (AUS) and SNR were determined for spectra of tumour samples mounted on SS and CaF_2 slides recorded with equal exposure times. Both AUS and SNR were found to be substantially greater for samples mounted on SS compared with those mounted on CaF_2 (Figure 2.6) the explanation of this is explored in the discussion.

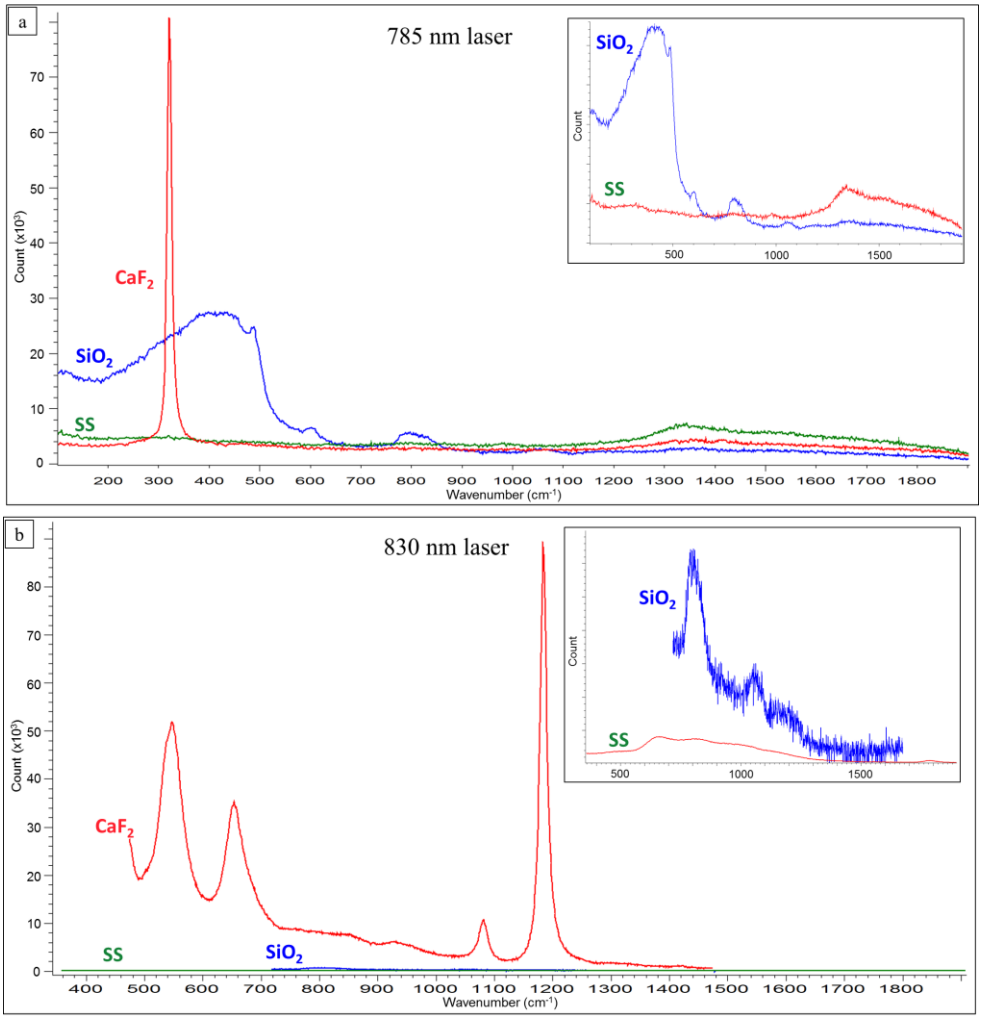
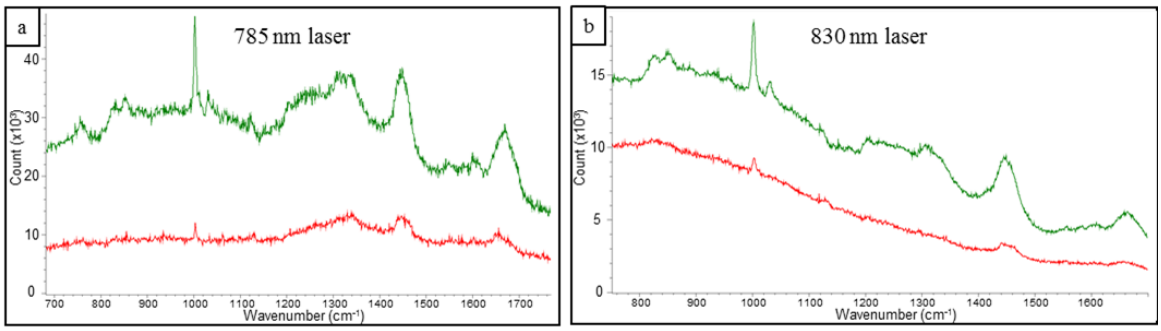


Figure 2.5 – Substrate background spectra. Background spectra of calcium fluoride (CaF_2 (red)), stainless steel (SS (green), fused silica (SiO_2 (blue)) using 830 nm and 785 nm lasers. Inset showing SiO_2 and SS only. Same exposure times used for each substrate.



SS	SNR= 21.2	AUS= 2,649,376
CaF	SNR = 9.1	AUS= 886,503

SS	SNR= 23.5	AUS= 8,916,931
CaF	SNR = 6.3	AUS= 4,758,422

Figure 2.6- Characteristic spectra of IDH-MUT tumour mounted on CaF_2 (red) and SS (green) slides using785 and 830 nm excitation lasers. SNR and area under spectrum (AUS) for each substrate with each laser wavelength is shown in table

2.3.4 Laser selection and effect on spectral characteristics

It has already been shown in Section 2.3.1 that use of the 785 nm laser resolves a larger number of Raman peaks in the raw spectrum when compared with the 830 nm laser. Preliminary data from the *Rapid molecular genetic classification of gliomas using Raman spectroscopy* study were acquired using both 785 nm and 830 nm lasers and these data were used to further evaluate the optimal laser for use in future Raman studies. In the preliminary dataset there were 15 WHO IV astrocytomas with IDH mutations (IDH-MUT) and 15 WHO IV astrocytomas without IDH mutations (IDH-WT). All samples were FFPE sections. A comparison of the mean SNR for all the IDH-MUT and IDH-WT spectra (32,524 and 30,877 spectra, respectively) is shown in Figure 2.7. SNR was calculated using the parameters outlined in

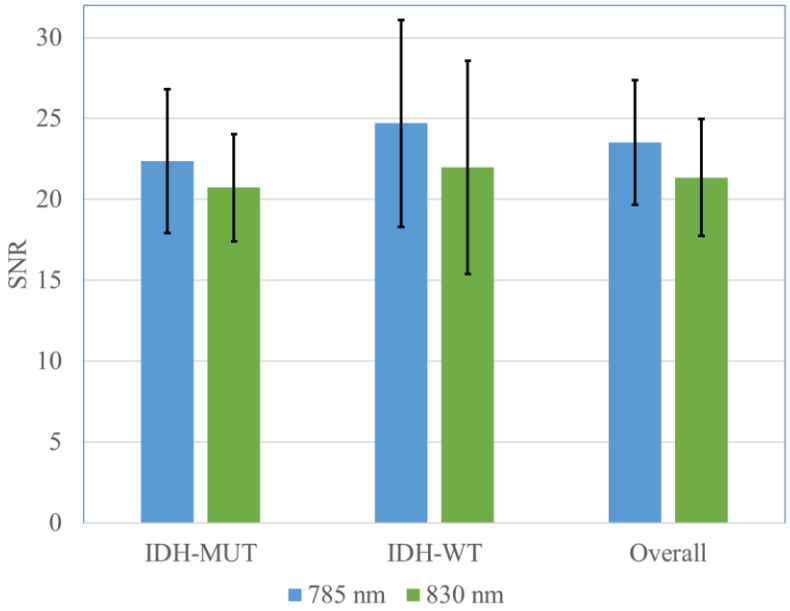


Figure 2.7 – SNR comparison of different lasers. Mean SNR of spectra for IDH-MUT and IDH-WT samples using 785 nm and 830 nm lasers. Overall mean is also represented. Error bars= 95% confidence intervals.

Section 2.2.4 and the same preprocessing steps were used for all the spectra as outlines in Chapter 4. There is found to be a slight SNR advantage when using 785 nm compared with 830 nm, although this was not statistically significant indicated by the broad and overlapping 95% confidence intervals. The size of the difference in SNR correlates well with the principle that Raman signal intensity is inversely proportional to the laser excitation wavelength[234,235] although the difference will also be a function of variations in background signal at the different wavelengths.

The preliminary data from the genetic subclassification of glioma study were subjected to supervised machine learning classification modelling (using Principal Component Analysis fed Linear Discriminant Analysis (PC-LDA)) and cross validation. The details of this analysis will be extensively covered in Chapter 4. For the purpose of laser comparison, the classification model performance using data collected using the 785 nm laser was compared with the performance using the 830 nm laser and demonstrated a small, but not statistically significant, improvement in performance when employing the 785 compared to the 830 nm laser (Table 2.2).

Laser used for measurements	**Sensitivity** (95%CI)	**Specificity** (95%CI)	**Accuracy** (95%CI)
785 nm	0.89 (0.82 to 0.94)	0.90 (0.88 to 0.93)	0.95 (0.84 to 0.99)
830 nm	0.86 (0.72 to 0.88)	0.88 (0.69 to 0.92)	0.90 (0.81 to 0.95)

Table 2.1 - Classification model performance using data collected with 785 nm or 830 nm lasers

2.4 Conclusions

This preliminary study allowed for sample and spectrometer optimisation. It was also used to test and evaluate the tissue workflow from brain bank to laboratory. The study findings are summarised in Table 2.2. With regards to pathology type, it was possible to acquire good quality spectra (SNR >10) of normal brain, both IDH-MUT and IDH-WT, and necrosis. This was true for both cryosections and FFPE sections. Wax artefacts were seen with some FFPE sections mounted on CaF_2 slides, but not with SS and SiO_2. This is consistent with published findings showing that dewaxing processes are less effective on CaF_2 slides[192,193]. The best substrate with these brain tissue samples was found to be stainless steel, which yielded a SNR of between 2 and 4 times higher than CaF_2. The use of mirrored stainless steel substrate

has been described by others, who reported a 1.43 increase in SNR using SS when compared with CaF_2[236]. They suggest that the increase in SNR is mainly due to the double pass of photons through the sample as a result of the reflection of the incident laser back through the sample as a result of the mirrored surface. In addition to this forward scattered Raman photons are reflected back into the spectrometer for the mirror surface. The increased AUC seen in the data presented here further supports this theory. The AUC reflects the intensity of the spectra and is interestingly also approximately double for samples mounted on SS relative to those mounted on CaF_2. The high cost of CaF_2 slides (~£100 per CaF_2 slide compared with ~£1 per SS slide) also makes them a less practical choice when large sample numbers are likely to be used in a study, and in the context of commercialisation of any technique at a later stage. Other issues with the CaF_2 slides include the difficulties with retained wax as mentioned previously, and the brittleness of the slides resulting in shattering if not handled with care. Moreover, there was poor sample adherence to the CaF_2 slides and the sections were often partially washed off in the dewaxing process. Sample adherence was, to a lesser extent, also a problem with SS slides. To address this issue a slide coating protocol was developed (Appendix A). The effect of this coating on sample adherence was investigated by another researcher in our group[viii] who demonstrated that on average 99% of the sample remained adhered to the coated SS slide, compared with 85% on uncoated SS, and an average of 80% of the sample remained adhered to the coated CaF_2 slide, compared with 40% on uncoated CaF_2.

[viii] Work undertaken by Natalie Pospisilova as part of Final Honors School research project (unpublished)

Variable	Summary of findings	Conclusion
Pathology type • Normal brain • IDH-WT • IDH-MUT • Tumour necrosis	• Raman peaks were resolved, even using raw spectra, with SNR >10 with all laser configurations tested on all pathology types	All pathology types suitable for sampling
Tissue preparation • Cryosection • FFPE section	• Both cryosection and FFPE sections give Raman spectra with SNR>10 • Cryosection and FFPE section spectra have similar Raman peaks • Cryosections are unprocessed compared to FFPE sections which are altered by fixation and embedding • Cryosections have a short life span compared to FFPE that can be reused • Risk of wax artefact with FFPE sections • Superior white light image with FFPE	Both cryosection and FFPE sections can be used successfully
Substrate • Calcium fluoride (CaF_2) • Stainless steel (SS) • Fused silica (SiO_2)	• SiO_2 contributes significant background • SS gave significantly better SNR than CaF_2 • SS significantly cheaper than CaF_2 • CaF_2 very brittle • Poor sample adherence to CaF_2 slide despite slide coating	SS slides provide the best overall spectra and are a better practical solution
Excitation laser • 830 nm • 785 nm	• Slightly better SNR with 785 nm (no statistical significance) • Better performance in classification model with 785 nm (no statistical significance)	Possibly small advantage of 785 nm laser

Table 2.2 - Summary of preliminary study findings of sample preparation and spectrometer configuration

With regards to excitation laser selection there was found to be a slight advantage, in terms of SNR and classification model performance, when using the 785 nm laser rather than the 830 nm. The optimum wavelength for biological sample Raman measurements has been extensively investigated by others[227].

Near-infrared lasers are generally considered most suited to Raman measurements of biological samples as they avoid the main tissue fluorescence background which peaks around 600 nm and tapers off around 750 nm[164]. An excitation wavelength of 785 nm seems to be favoured by most groups undertaking Raman measurements on biological tissue[182,184,190,194,236]. The data presented here also found slightly better performance for 785 nm out of the two wavelengths investigated. This was also favourable from a more practical point of view, in terms of the available Raman equipment. Renishaw, the industry collaborators, provided all of the spectroscopic equipment for this project, and manufacture two types of Raman microscope system:

1. **inVia™**: a research-grade system that provides access to, and options to change, all internal optics, as well as options for multiple excitation lasers. The system has a large footprint and is a Class 3B[ix] device requiring a door interlocking system to comply with laser safety regulations;
2. **RA800-series**: is a compact bench-top system with fixed optics (785 nm excitation laser) and inbuilt self-calibration capabilities. The system is a Class 1[ix] laser product requiring no additional laser safety precautions. The system is designed for clinical application and has a simple user interface (see Chapter 4 for further details).

The preliminary work described in this Chapter was performed on the inVia™ system in order to allow for testing of various spectrometer parameters. However, given that an important aim of this project is to enable the translation of Raman tissue analysis into neuropathology and neurosurgery clinical settings, the RA800-series benchtop spectrometer was the preferred choice for loan and use in subsequent measurements. The bench-top spectrometer was thus installed in a purpose-designed room (requiring the installation of anti-vibration benches) in the Department of Neuropathology which is near to the neurosurgical operating theatres. Having the spectrometer on site not only allowed for faster throughput of brain bank cryosection and FFPE samples but, more importantly, allowed for the analysis of fresh tissue samples straight from the operating theatre. This was vital to allow Raman spectroscopy to be integrated into the neurosurgery clinical workflow.

[ix]Laser Classes are defined by the International Electrotechnical Commission standard IEC60825-1. The laser class reflects the risk of injury from laser exposure. Class 1 represents no risk under conditions of normal instrument use. Class 3B lasers are hazardous if the eye is exposed directly.

Chapter 3. Experimental parameters and effect of laser exposure on Raman spectra

A preliminary investigation into the genetic subclassification of gliomas by Raman spectroscopy was performed in order to demonstrate proof of concept and to refine and identify any problems with the experimental methods. The use of these preliminary data has already been reported in Section 2.3.4, to compare the performance of different excitation lasers. This chapter outlines the findings of this preliminary investigation and reports on experiments to evaluate specific variables which may affect the Raman spectra. Throughout these experiments the spectral changes due to increased laser exposure, including suppression of background signal, are characterised, and a method is proposed for reaching a compromise between good quality spectra and limited spectral change due to laser exposure.

3.1 Preliminary experiment: 4 second exposure versus 6 second exposure time

As outlined in Chapter 2, the preliminary investigation into the genetic subclassification of gliomas by Raman spectroscopy comprised of 15 WHO IV astrocytomas with IDH mutations (IDH-MUT) and 15 WHO IV astrocytomas without IDH mutations (IDH-WT). Samples were dewaxed and mounted on coated stainless-steel slides. Raman measurements were made using the inVia™ spectrometer, with an excitation wavelength of 785 nm. For this preliminary investigation, two spectral maps were collected per sample from different locations in areas identified as tumour by the neuropathologist on adjacent H+E slides. The maps were acquired using Renishaw's StreamLine™ imaging protocol, which provides rapid sample imaging by scanning with a line laser and synchronising the movement of the laser line on the sample surface with the CCD detector. This scanning method was chosen to maximise the area of the sample covered by the laser whilst minimising the overall map collection time. Each map corresponded to an area of 1000 μm^2 on the sample surface and a total of 2025 spectra were collected per map with a step size along both the x and y axes of 22.7 µm[i]. The exposure time was set to 4 seconds per spectrum. After collection maps were pre-processed to remove background and cosmic rays, and a Principal

[i] The step size chosen corresponded to a Y-bin of 16 within the acquisition software which sums together the signal from 16 pixels on the CCD. This increases the signal to noise ratio but at the expense of spatial resolution which is justified in this application as these maps were not indented to be used for image generation.

Component-Linear Discriminant Analysis (PC-LDA) classification model was used to attempt to classify the spectra into IDH-MUT and IDH-WT groups. Details of the spectral pre-processing and PC-LDA modelling will be provided in Chapter 4.

After 15 samples (7 IDH-MUT, 8 IDH-WT) had been imaged, the data were reviewed in order to establish spectral quality and the performance of the PC-LDA classification model. It was noted that although a significant number of spectra showed an acceptable signal to noise ratio, a large number showed very little signal at all, instead yielding only a large background signal which swamped the Raman signal. The percentage of spectra per map that showed only background ranged from 5% to up to 80% in some maps. To identify the origin of the background, samples were reviewed by the neuropathologist to look for any common histological features in the samples with strong background, but none could be found. Both IDH-MUT and IDH-WT samples had similar number of spectra with a dominant background. The background due to the objective lens and the substrate were also characterised and differed from the dominant background spectra in both intensity or spectral shape and thus were not thought to be the cause of the background. The literature reports on the problem of a broad background in Raman spectra of biological tissue and it is a well-recognised phenomenon[237,238]. The origin of the background is under debate with some contesting the generally accepted view that it is due to sample fluorescence[239,240]. This will be explored further at the end of this chapter.

To filter out spectra with weak Raman signal due to a dominant background, a signal to noise ratio (SNR) threshold filter (removing all spectra with SNR<4) was incorporated into the pre-processing steps prior to building the classification model[ii]. In addition, the exposure time was increased to 6 s per spectrum for the remaining samples in the preliminary study (21 further cases). Although this would only have a small effect on SNR, increasing exposure time further resulted in frequent CCD saturation.

The preliminary data therefore consisted of 15 samples recorded with a 4 s exposure time and 21 samples recorded with a 6 s exposure time. There was a 12% reduction in the number of spectra with dominant background when using 6 s compared to 4 s exposure and a small increase in overall SNR. Despite this,

ii The number of spectra removed by any pre-processing step was closely monitored to prevent over processing of the data and therefore overfitting of a model. This is addressed in more detail in Chapter 4 but, in general, it was considered acceptable to remove up to 30% of spectra in the preprocessing stage and the SNR filter threshold was set accordingly.

comparison of classification model performance showed much better performance for the 4 s data (0.92 sensitivity, 0.87 specificity) compared with the 6 s data (0.69 sensitivity, 0.39 specificity) when employing equivalent pre-processing steps. The mean spectra, after pre-processing, for the IDH-MUT and the IDH-WT groups and the difference spectra[iii] (IDH-MUT minus IDH-WT) for the 4 second and the 6 second data sets are shown in Figure 3.1. The 4 s and 6 s mean spectra have an intensity difference that is in keeping with the 2 s difference in exposure time (Figure 3.1a). The difference spectra show a nearly complete reversal of the prominent peaks when moving from a 4 s to a 6 s exposure time (Figure 3.1b-d). Figure 3.1d overlays the 4 s and 6 s difference spectra with the variation between samples shown by the shaded regions representing one standard deviation. Even allowing for the relatively large variance between samples (which is not unexpected in heterogeneous biological tissue), prominent peak differences persist.

To further investigate the difference observed between the 4 s and 6 s data the PC-LDA model was interrogated. As part of the model building, Principal Component Analysis was used to reduce the data dimensionality, with analysis of variance (ANOVA) used to select only those principle components (PCs) that describe a statistically significant difference in the data ($p<0.001$). The first principle component, PC 1, was the most significant component (highest ANOVA F-score) in both the 4 s and 6 s models. The PC 1 loadings are plotted in Figure 3.2, and show a clear difference between the 4 s and 6 s models which mirror the changes seen in the difference spectra in Figure 3.1. The models were run several times which consistently reproduced this difference (it is accepted that PCA is capable of reversing loadings between models).

[iii] Difference spectra are calculated by subtracting the normalised mean spectra of each group.

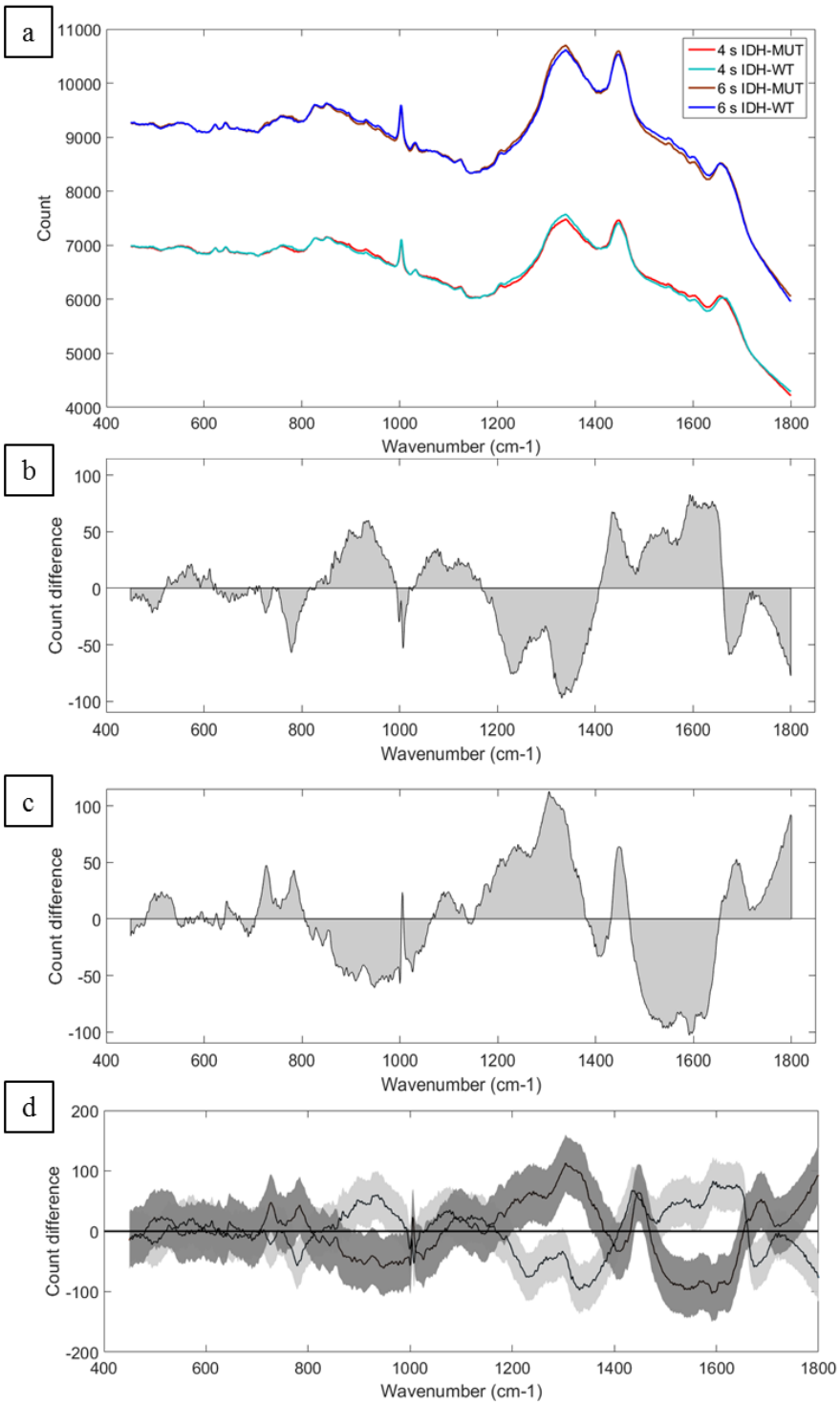


Figure 3.1 – Preliminary 4 s and 6 s data. a) Normalised mean spectra for IDH-MUT and IDH-WT samples in the 4 s and 6 s model. b) IDH-MUT minus IDH-WT difference spectra for the 4 s model. c) IDH-MUT minus IDH-WT difference spectra for the 6 s model. d) Combined 4 s and 6 s IDH-MUT minus IDH-WT difference spectra with shaded area giving 1 standard deviation around mean.

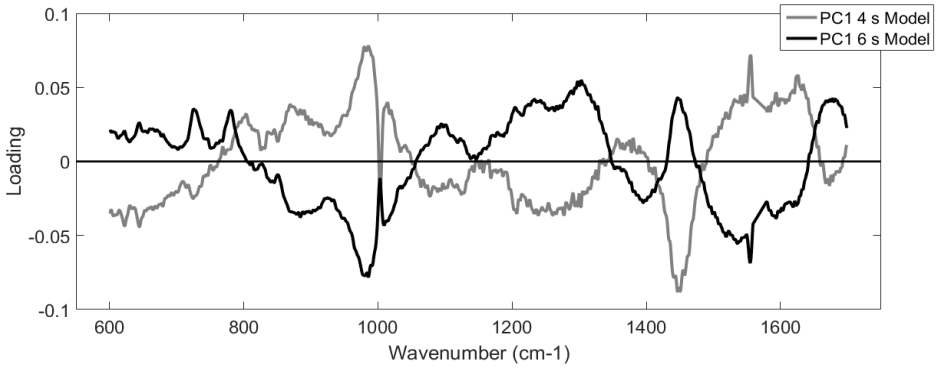


Figure 3.2 -Principle component 1 loading plot for 4 s and 6 s models

Several factors were considered in order to explain the difference between the 4 s and 6 s data sets:

1. **Pre-processing bias** – The pre-processing steps were reviewed to look for a possible bias in performance towards the 4 s data. Minimal pre-processing was performed on the spectra, but several combinations and changes in parameters (e.g. polynomial order used for baseline correction) were trialed and showed no evidence of 4 s data bias. Indeed, even with the raw data the performance of the 4 s data classification model was better than that of the 6 s model.
2. **Number of samples** – The 6 s data set contained more samples than the 4 s data, which may have resulted in a larger intergroup variation and therefore worse model performance. Multiple random selections of equal numbers of samples from each data set showed no change in the relative performance of the two models, ruling out this possibility.
3. **Sample inherent heterogeneity** – Could the samples used in the measurements employing a 6 s exposure time be inherently more heterogeneous than those used with a 4 second exposure time, leading to worse performance of the classification model for the data set with the longer exposure time? To investigate this, the samples employed for the 6 s measurements were subjected to a repeated set of measurements, using a 4 s exposure time. Due to limited availability of FFPE tissue, the same sections were used but at different locations on the sample from the original maps. The mean spectra, difference spectra and PC 1 loadings were all similar to those obtained for the original 4 s data sets. The classification model performance was also improved compared with the original 6 second data (sensitivity 0.84, specificity 0.72). This would suggest that there is no difference in the quality or heterogeneity of the samples used in the original 6 s exposures, as improved classification was possible when the experiment was repeated with a 4

s exposure. One confounding factor here was the fact that these samples had previously been exposed to the laser, albeit at a different location. To address this issue an experiment was undertaken to *investigate the effect of laser exposure on surrounding tissue.* This study will be presented in Section 3.2.

4. **Degradation in samples over time** - All samples were sectioned and dewaxed in one batch and then the data collected over a period of weeks. The 4 s data was collected between 1 and 3 weeks after sample preparation and the 6 s data between 4 and 6 weeks. Could there be a degradation in sample over time which could explains the difference between the two datasets? To answer this question an experiment was undertaken to *investigate the effect of sample degradation over time*. This study is reported in Section 3.3.
5. **Change in spectra due to laser exposure** – The preliminary experimental data raise the possibility that increasing the exposure time by 2 s may cause a change in the spectra which results in a change in the classification model performance. This hypothesis was supported by the observation in the preliminary experiment of substantial suppression of the background intensity with repeat laser exposure over the same tissue location (see Figure 3.12 for examples of this). Characterisation of this background suppression and investigation of the possibility of laser-induced spectral changes was investigated as part of the previously mentioned experiments.

The experiments to investigate the effect of laser exposure on surrounding tissue and the effect of sample degradation over time will now be reported. The main aim of each experiment was to answer the questions raised above, and to relate the results back to the observed discrepancy between the data sets recorded with 4 s and 6 s exposure times. An additional aim was to investigate the effect of increasing laser exposure on both the background suppression and possible changes induced in the spectra.

3.2 Investigation into the effect of laser exposure on surrounding tissue

The main aim of this experiment was to establish if there was an effect on the spectra due to previous laser exposure of an adjacent area on the same sample. Additional aims were to:

1. establish if previous laser exposure may be a confounding factor when repeating the Raman measurements of 6 s samples with a 4 s exposure time;

2. determine whether FFPE sections can be used multiple times, with different map locations, safe in the knowledge that there should be no change in the sample induced by previous adjacent laser exposure (thus allowing full utilisation of the valuable and limited FFPE tissue resource);
3. if 2 is not the case, make inferences about the effect on a localised area of a sample when the adjacent tissue is irradiated by the laser.

3.2.1 Method and setup

Two FFPE tumour samples, one IDH-MUT and one IDH-WT, both with large areas of heterogeneous tumour, were selected for this experiment, with measurements carried out on the inVia™ spectrometer. A 1 mm^2 map, as used for the preliminary experiments described in Section 3.1, was recorded with a 6 s

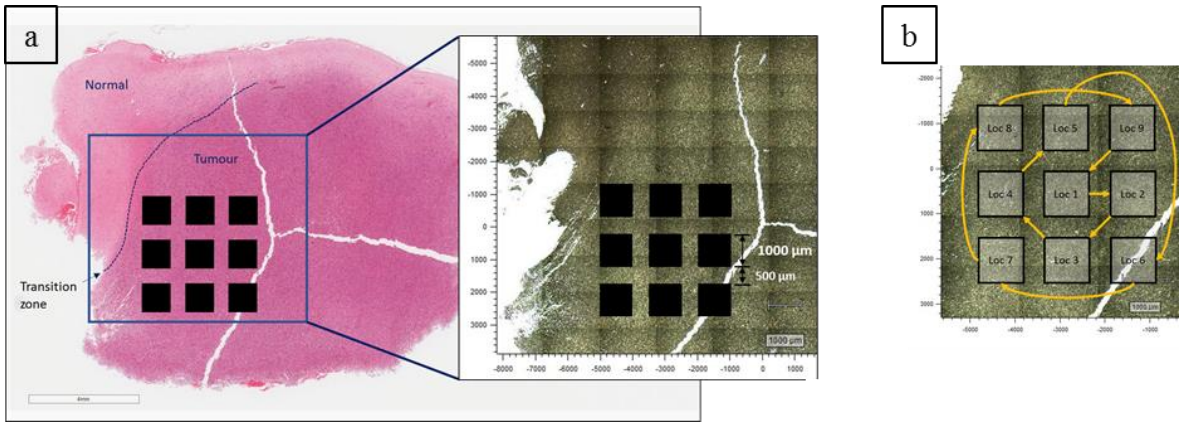


Figure 3.3 - Experimental setup for laser exposure on adjacent tissue. a) Map layout and location on IDH-mutant sample. b) Sequence of map measurements from location 1 to 9 and then completing with a repeat map over location 1.

exposure time (location 1). A further eight maps (location 2 to 9) were then measured in eight further locations adjacent to location 1 as indicated in Figure 3.3. After the map from location 9 was acquired, a repeat measurement at location 1 was undertaken to provide a baseline for the change in spectra due to repeat laser exposure (including any background suppression due to repeat exposure). The sequence of map measurements is outlined in Figure 3.3b, with the location number corresponding to the sequence of collection. The laser line was measured to have a dimension of 67 µm in the horizontal plane, and therefore a space of 500 µm was left in all directions between adjacent maps so that there was no risk of overlap of direct laser exposure between maps.

Pre-processing steps included spatial median filtering to remove cosmic ray artefacts and normalisation using Extended Multiplicative Scatter Correction (EMSC) including a third-order polynomial baseline.

Maps were compared using mean spectra, change in overall spectral intensity, change in SNR and differences in classification model performance. For intensity comparisons the spectra were compared without normalisation. Full explanations of EMSC and SNR calculations can be found in Chapter 4, Section 4.2.5. Box and whisker plots were used to describe the data graphically. Location 6 in the IDH-MUT sample spanned a crack in the tissue, which needed to be taken into account in the data analysis.

3.2.2 Results

Both the IDH-MUT and the IDH-WT samples gave very similar results, and so for the sake of simplicity, only the results from the IDH-MUT samples are reported here. Figure 3.4 illustrates the mean spectra

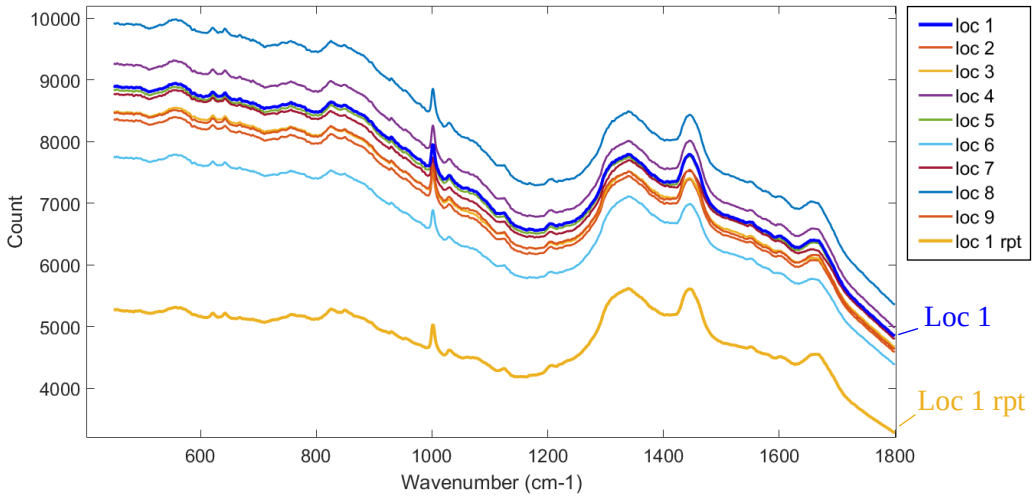


Figure 3.4 – Effect of laser exposure on adjacent tissue. Mean spectra from each map location 1 to 9 and then location 1 repeat (Loc 1 rpt).

for each map location. The background suppression effect due to previous exposure is clearly seen in the repeat measurements on location 1. This effect is not observed in any of the other adjacent maps, which would suggest they are not affected significantly by adjacent laser exposure. The spectral peaks are also very similar in every location. Figure 3.5 shows the distribution of the map intensities (at phenylalanine peak, 1001 cm^{-1}) and the SNR and confirms that the maps are all within the same range of intensities and SNR, allowing for some variation as would be expected in heterogeneous biological samples. The intensity range is reduced in the repeat measurement of location 1, as a result of the background suppression referred to above. There is also an increase in the SNR range in this location.

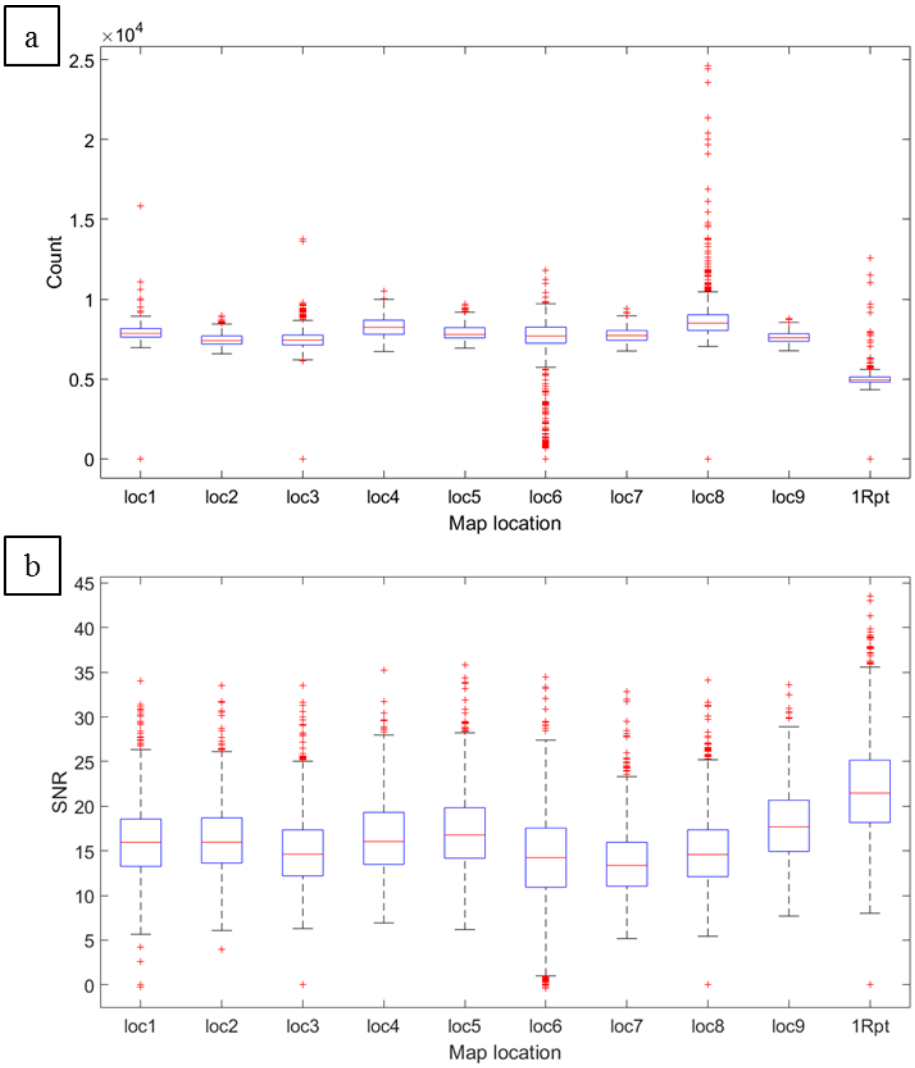


Figure 3.5 – Box and whisker plot of adjacent maps. a) Intensity at 1001 cm^{-1} and. b) SNR for each map location. Box and whiskers show 25%-75% and 1-99% percentile, respectively. The median is also indicated. (1Rpt=Repeat measurement of location 1)

The spectra from locations 2 and 4 were divided into areas immediately adjacent ("near") to and areas relatively far away ("far") from the previously exposed tissue in location 1, as illustrated in Figure 3.6. Each area included 250 spectra. These were reviewed separately to see if there was any evidence of background suppression in the spectra recorded for areas immediately adjacent to the exposed tissue compared to spectra far away from exposed areas. Figure 3.6 shows the mean spectra for "near" and "far" areas of location 2 and 4. For comparison the mean spectrum and standard deviation of the whole map in that location are shown. There was no consistent decrease observed in the intensity of the "near" spectra compared to the "far" spectra, and the locations and intensities of spectral peaks were similar in every case.

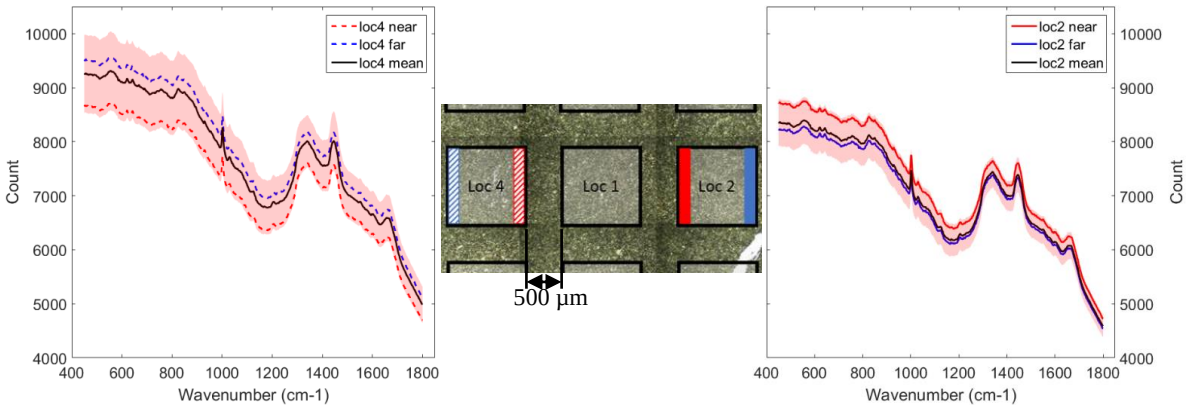
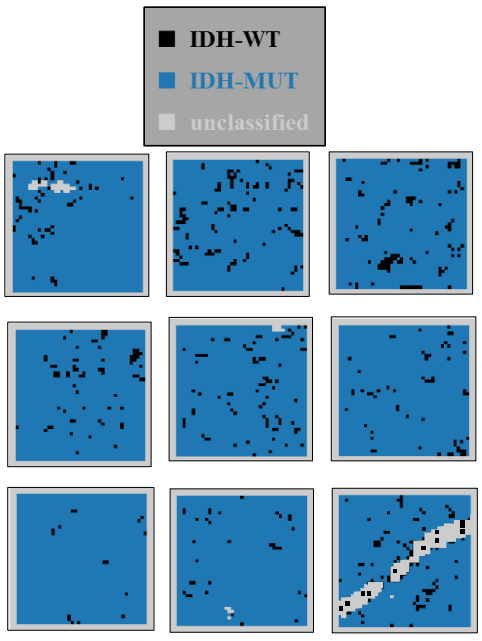


Figure 3.6 – Graphical illustration of "near" (red) and "far" (blue) regions, relative to location 1, of location 2 (solid fill) and location 4 (hash fill). Corresponding mean spectra for "near" and "far" regions with the mean (black) and standard deviation (shaded area) of the whole map shown for reference.

The spectra from each map location were classified into IDH-WT or IDH-MUT groups using the original classification model built using the 6 s data. Figure 3.7 shows that the percentages of spectra correctly classified are similar in all locations except for location 6 which has more unclassified spectra due to the crack in the tissue[iv]. The false colour map confirms that there was no correlation between classification group and proximity to previously exposed areas.



Location	% classified IDH-WT	% classified IDH-MUT	% unclassified
Loc8	3	78	19
Loc5	6	77	17
Loc9	6	77	17
Loc4	4	80	17
Loc 1	4	79	17
Loc 2	3	80	17
Loc7	1	82	17
Loc3	2	81	17
Loc 6	13	55	32

Figure 3.7 - False colour map and corresponding percentage of spectra classified into IDH-wildtype, IDH-mutant and unclassified using original 6 s classification model.

[iv] All map locations have spectra that are unclassified around the perimeter of the map as seen in Figure 3.7. This is because spatial median filtering (to remove cosmic ray artefacts) is not fully reliable at the map edges and thus one spectrum around the perimeter is omitted from analysis.

3.2.3 Interpretation

This experiment suggests that there is no effect of laser exposure on adjacent tissue in terms of mean spectral intensities, spectral peak characteristics, SNR or IDH classification model performance. Adjacent laser exposure is not a confounding factor when repeating the Raman measurements of 6 s samples with a 4 s exposure time and therefore the repeated data reported in Section 3.1 are valid, supporting the notion that there is no difference between the heterogeneity of the samples used for the original 6 s and 4 s exposures.

The background suppression effect observed on repeated laser exposure of the same area does not appear to affect surrounding areas. Even spectra measured only 500 µm away from previously exposed area show no significant background suppression. These data confirm that FFPE sections can be used repeatedly, as long as the map location is different, with no concern about effects from previous laser exposure.

3.3 Investigation into sample degradation over time

The main aim of this experiment was to investigate the effect on the Raman spectra of increasing the time interval between sample preparation and measurements. Additional aims were to:

1. investigate the background suppression with repeated laser exposure and its effect on signal intensity, SNR and IDH-MUT and IDH-WT mean and difference spectra ("Consecutive maps");
2. establish if there is any recovery in the background suppression over time ("Cumulative maps").

3.3.1 Method

Three FFPE samples were selected that had large homogenous areas of tumour (one IDH-WT and one IDH-MUT) or normal white matter (one normal cortical brain). Eight parallel sections were cut from each sample and prepared for Raman measurement on the same day. Preparation protocols, Raman map dimensions and spectral pre-processing were the same as for the previous experiment (Section 3.2.1). The exposure time was 4 s per spectrum[v]. Between measurements all samples were stored in a slide box at room temperature.

[v] An exposure time of 4 s was chosen as a result of the initial 4 s IDH preliminary experiment described in Section 3.1, in case the 6 s acquisition had caused a change in the spectra.

There was a requirement in this experiment to match each section orientation to ensure that the same locations were mapped each time on the new section, and that repeated maps at the same location on a section could be recorded once the slide had been removed and then reinserted into the spectrometer. To achieve this the slides were placed in a rotational slide holder that allowed for subtle adjustments in the slide orientation. Samples were orientated using a two-point correlation, in which two distinctive features were located on the section (Figure 3.8). The first feature was set as the zero-reference point for the coordinates and then the slide was orientated so that the co-ordinates of the second distinctive feature

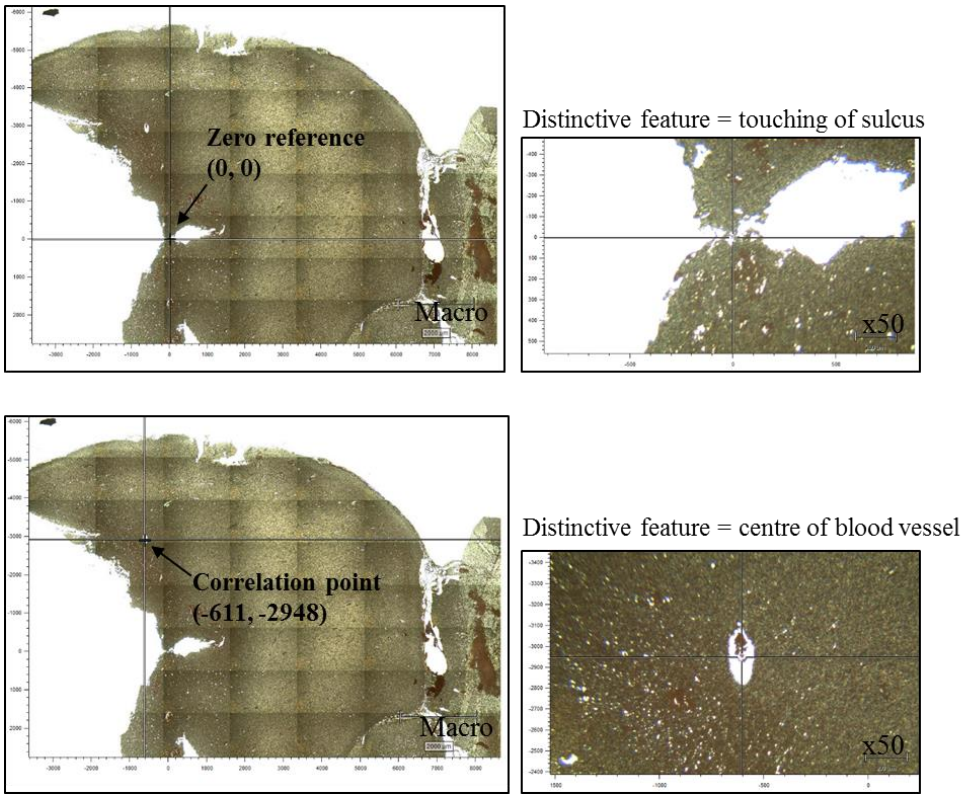


Figure 3.8 - Section orientation method with distinctive features chosen, touching of sulci and centre of blood vessel here, for zero x-y reference and correlation point (normal brain sample shown here).

(the correlation point) were the same on each section. Due to the slight variation in the parallel sections this approximate orientation technique ensured a satisfactory correlation of map locations between sections and a more accurate correlation when aligning the same section for repeated sampling as the distinctive features were the same.

After sample cutting and preparation (on week 0 day 0), measurements were made on week 0 day 1 and week 0 day 2 and then every week for 5 weeks (week 1 to 5). For each of the three sample types (IDH-

MUT, IDH-WT and normal brain) the following combinations of section measurements were performed (Figure 3.9):

- **Main experiment** – A new section (section 1 to section 7) was measured at each time point at the same location each time. This was to address the main experimental question of whether there is a change in spectra with time after sample preparation.
- **Consecutive maps and background suppression recovery** - Seven consecutive maps were collected on the same section (section 8) on the same day (week 0 day 1) and then 2 further maps were collected on week 5. All maps were recorded at the same location. This was done to investigate the background suppression effect and to assess if there is recovery in the suppression effect after 5 weeks.
- **Cumulative maps** – Repeated maps were carried out on the same location on the same section (section 1) at each time point, giving a total of seven overlapping maps. This was to investigate if the background suppression effect was present and persisted in the spectra over the time course.

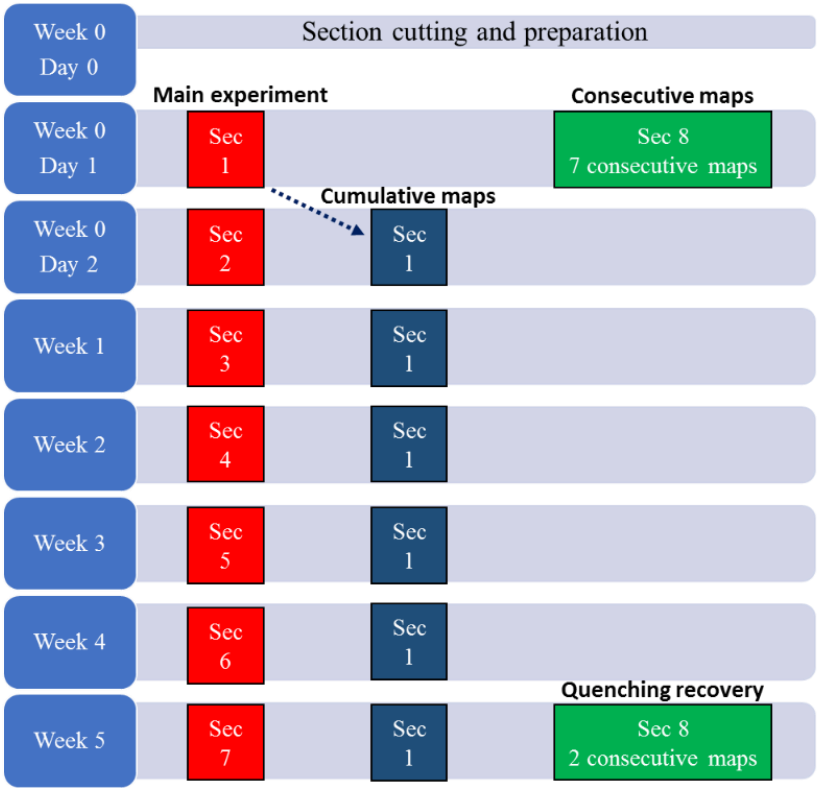


Figure 3.9 - Experimental timeline for sample degradation over time experiment. Main experiment (red), cumulative maps series (dark blue), consecutive maps and quenching recovery series (green). Sec=section

For the purpose of intensity comparison, mean spectra from each map were compared, with no normalisation performed in the pre-processing. To compare IDH-MUT versus IDH-WT spectra, mean and difference spectra were analysed and EMSC normalisation was used.

3.3.2 Results

There was no evidence of sample degradation over the 5-week time course in either of the tumour samples or the normal cortical brain sample. Figure 3.10a shows the mean spectra at each time interval for the IDH-MUT sample. The IDH-WT and normal brain showed similar mean spectra to this. The intensity (at phenylalanine peak, 1001 cm^{-1}) and the SNR did fluctuate over the time course, but this was acceptable within the normal fluctuations seen in the heterogenous parallel sections and there was no temporal correlation to the fluctuations (Figure 3.10b and c). The Raman peak positions and intensities remained constant throughout the five weeks.

The mean spectra of maps taken on week 0, week 3 and week 5 were compared for both the IDH-MUT and IDH-WT samples and are illustrated in Figure 3.11. There was no change in the difference between the two tumour samples over the time course, with the variation in the difference spectra at week 3 and week 5 falling within one standard deviation of the difference spectrum at week 0 (Figure 3.11b). The spectra from each map taken at each time point were classified into IDH-MUT and IDH-WT groups using the original 4 s IDH model. Similar classification performance was seen throughout the time course, with a mean percentage of spectra correctly classified in each map of 75% (range: 66 – 80%) IDH-MUT spectra and 71% (range: 69 – 81%) IDH-WT spectra.

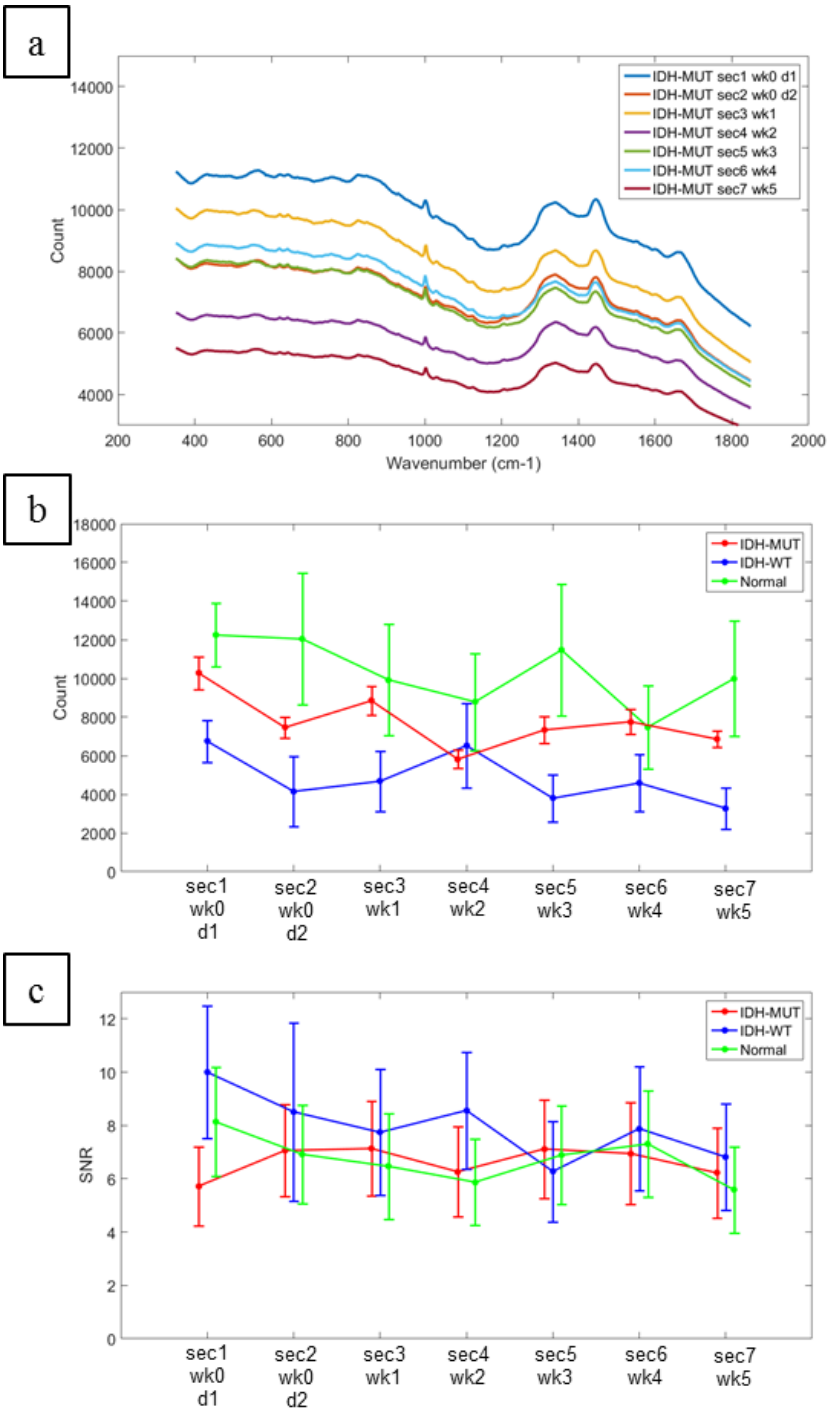


Figure 3.10 – Degradation over time. a) Mean spectra at each time interval from week 0 day 1 to 5 weeks after sample preparation in IDH-MUT tumour. b) Mean intensity at 1001 cm^{-1} and c) mean SNR over the time interval for IDH-MUT, IDH-WT and Normal samples

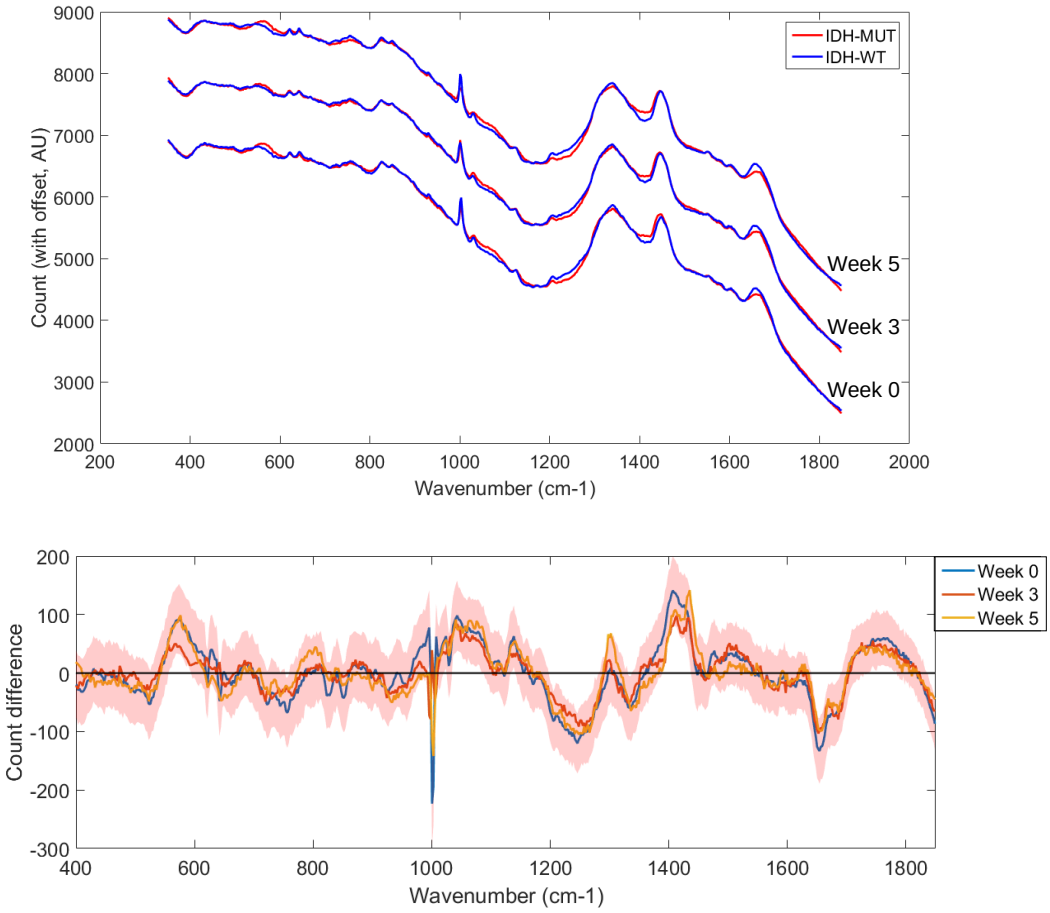


Figure 3.11 – Difference in IDH-MUT and IDH-WT spectra at week 0, week 3 and week 5. a) Mean spectra from map of IDH- MUT and IDH-WT at each time point (normalised mean spectra; plotted with offset for clarity). b) Difference spectra of IDH-MUT minus IDH-WT at each time point (shaded area= standard deviation of mean spectrum at week 0).

The cumulative maps (repeated maps on same location at each time point) and the consecutive maps (seven maps collected on the same section on the same day) are shown in Figure 3.12a for the IDH-MUT tumour sample. IDH-WT tumour and normal brain samples demonstrated similar trends. The previously observed suppression in background intensity resulting from repeated laser exposure was seen with consecutive maps. When a further map was collected at the same location five weeks later, there appeared to be very little recovery in the background, and indeed the suppression continued with the second repeated map (Figure 3.12c). A similar intensity suppression was seen for the cumulative maps which further confirms that the background suppression does not recover over time (Figure 3.12b).

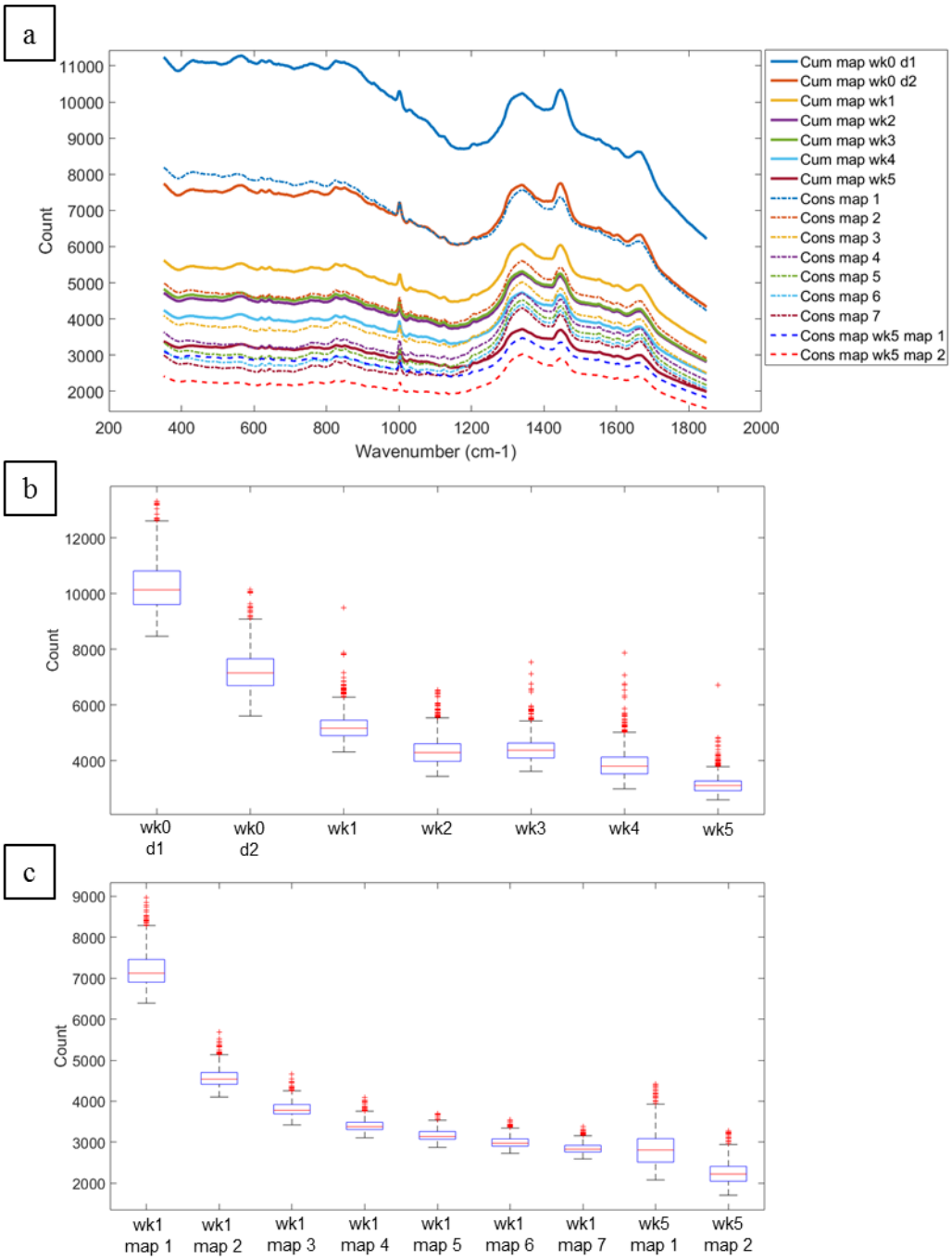


Figure 3.12 - Suppression of background intensity with repeated laser exposure. a) Mean spectra of all cumulative and consecutive maps. Box and whisker plot of intensity at 1001 cm^{-1} for b) cumulative maps over the time interval and c) consecutive maps in week 1 and then repeated in week 5. Box and whiskers show 25%-75% and 1-99% percentile, respectively. The median is also indicated.

The repeated consecutive maps of the IDH-MUT and IDH-WT samples are compared, there appears to be a slight decrease in the differences seen between the mean spectra of the two tumour types as the number of accumulations increases (Figure 3.13). Although the difference spectra of the 3rd, 5th and 7th repeated maps are within one standard deviation of the 1st map, there is a trend of decreasing difference. In contrast to this, the SNR increases as the number of consecutive maps increase (Figure 3.14), which

is likely related to the decrease in background noise. There is also better classification performance using the original 4 s IDH model, with the percentage of spectra in each map classified correctly increasing from a mean of 78% in map 1, up to 82% by map 3, and then reaching a plateau around map 4 of 84% correctly classified.

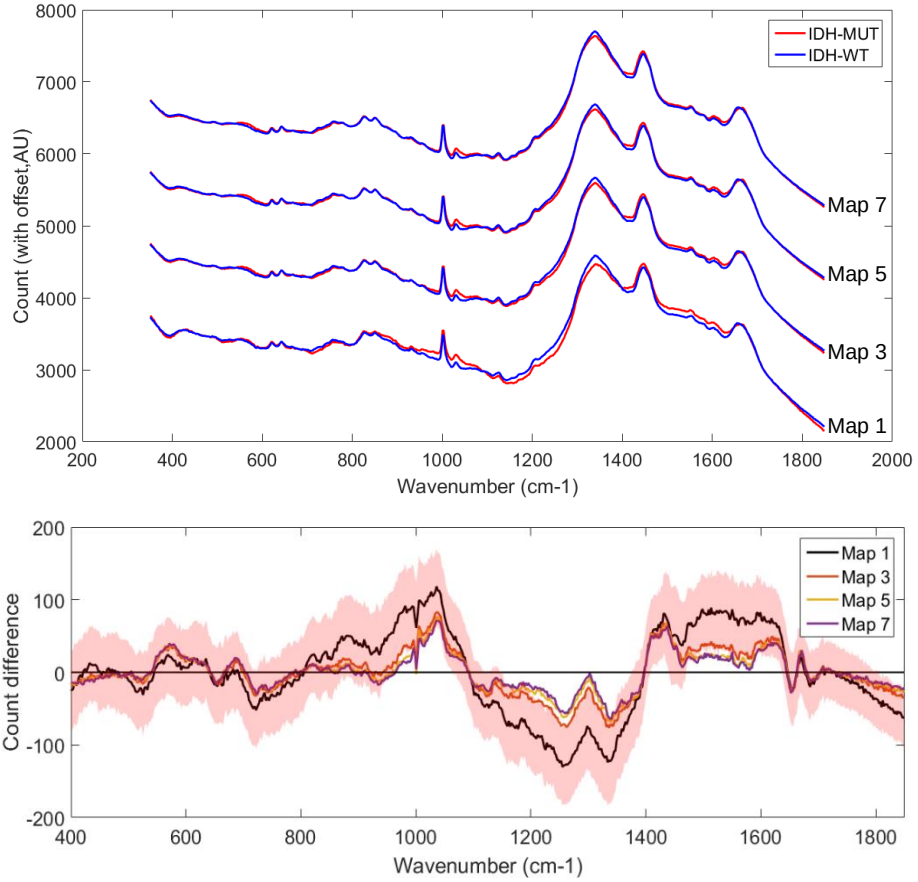


Figure 3.13 – Difference in IDH-MUT and IDH-WT mean spectra with consecutive maps. a) Mean spectra from map of IDH- MUT and IDH-WT at 1, 3, 5 and 7 repeated maps (normalised mean spectra; plotted with offset for clarity). b) Difference spectra of IDH-MUT minus IDH-WT at 1, 3, 5 and 7 repeated maps (shaded area= standard deviation of mean spectrum of map 1).

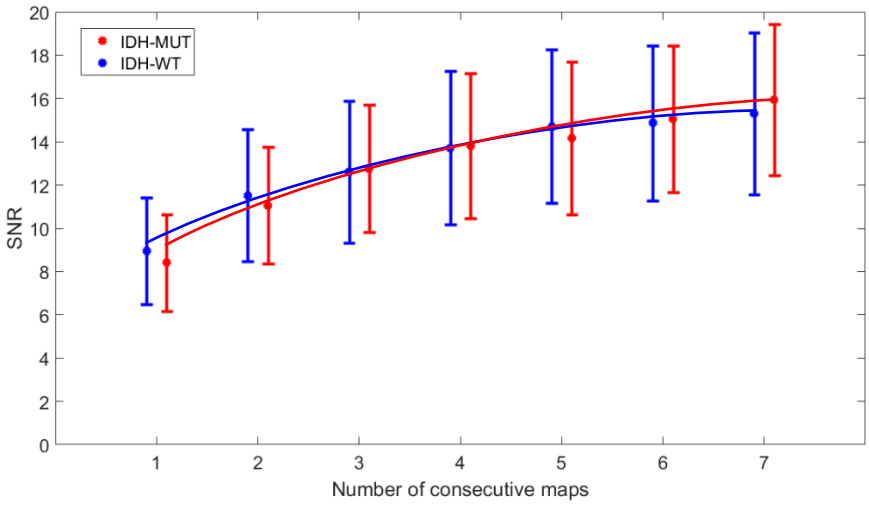


Figure 3.14 – Change in mean SNR per map with consecutive maps for IDH-MUT and IDH-WT samples (mean SNR with standard deviation error bars).

3.3.3 Interpretation

The key finding in this experiment is that there is no significant sample degradation over time, with Raman peak positions and intensities, and IDH-MUT minus IDH-WT difference spectra unchanged as a function of time from sample preparation to Raman measurement over a time course of five weeks. Degradation of samples over time cannot, therefore, explain the observed difference observed between the original 4 s and 6 s IDH data. Reports in the literature regarding FFPE sample degradation over time focus on antigen alterations or loss and relate to stored slides with glass coverslips or paraffin blocks and report some degradation over many years[241]. The tissue set up used here, dewaxed sections left to air dry on a slide with no coverslip, is unique to this particular application but these data agree with the general concept that FFPE samples are stable over time and additionally confirms that the Raman spectra remain unchanged.

This experiment also explored the background suppression seen with repeated laser exposure of a given location on a sample. The suppression is persistent over time, with no evidence of recovery. The effect on the IDH tumour classification model appears to be to increase performance with increased laser exposure. This correlates with a similar increase in SNR with repeated exposures, despite a trend towards a decrease in the magnitude of the IDH-MUT minus IDH-WT difference spectra.

Both this and the previous experiment have addressed, and ruled out, key variables which could potentially have caused the observed differences between the spectra recorded with 4 s and 6 s exposure times. The remaining hypothesis is that the extra laser irradiation time in the case of a 6 s exposure

induces spectral changes. This is supported by the evidence relating to the background suppression, which shows that in addition to the decrease in background there is also an increase in SNR and better classification model performance with repeated laser exposure. This cannot explain the poor performance of the 6 s model, as the evidence would suggest improvement in classification with the increased exposure time. However, there may be a point at which the effect of laser induced spectral changes outweighs the effect of improved SNR.

To inform the experimental design of the main glioma genetic classification study, it was important when choosing the exposure time to establish the optimum balance between preventing laser induced changes in the spectra and acquiring meaningful spectra with sufficient signal to noise. In an attempt to estimate the point at which laser induced changes outweigh any improvements to the SNR, an Allan variance analysis was carried out.

3.4 Allan variance

Allan variance, or two-sample variance, quantifies the point to point variability within a system[242]. The aim in this experiment was to use Allan variance to quantify the point to point variability within a sample Raman spectrum as a function of exposure time, and therefore to establish the optimum exposure time for use in future measurements. The Allan variance (σ_A^2) of *N* data points x_i collected during an exposure time t is defined as:

$$\sigma_A^2(t) = \frac{1}{N}\sum_i^N \left((x_{i+1}(t) - x_i(t)\right)^2 \quad [1]$$

At short exposure times, increasing the exposure time will usually lead to a reduction in the Allan variance as a result of reducing the measurement noise and thus improving SNR. However, at longer times, increasing the exposure time can lead to an increase in the Allan variance because of longer-term drift in experimental parameters. With reference to the brain tumour FFPE samples used in this study, there appears to be a change in the spectra when laser exposure times are increased, and it is hypothesised that this could result in an increase in the Allan variance. The point of minimum Allan variance would represent a compromise between variance due to low exposure time and variance due to changes in the tissue spectra, the latter of which may represent spectral changes due to laser exposure.

In addition to undertaking the Allan variance analysis, these data were also used to further characterise the rate at which the background suppression changes as a function of exposure time. Rates of intensity change in IDH-MUT and IDH-WT tumour samples were also compared.

3.4.1 Method

FFPE sections of two IDH-MUT and two IDH-WT tumour samples were measured on the RA802 spectrometer. The total exposure time was 10 seconds in the following combinations of individual times:

- 100 spectra at 0.1 s
- 50 spectra at 0.2 s
- 20 spectra at 0.5 s
- 10 spectra at 1.0 s
- 5 spectra at 2.0 s
- 2 spectra at 5.0 s

The location was changed between each 10 s exposure, and each of the above timing combinations was repeated in two different locations. The raw (unprocessed) data were used for variance calculations[vi]. The Allan variance was calculated using Equation [1] with the addition of a proportion step (dividing by $x_i(t)$), to give the relative change at each time point, using the following equation:

$$\sigma_A^2(t) = \frac{1}{N}\sum_i^N \left(\frac{(x_{i+1}(t) - x_i(t)}{x_i(t)}\right)^2 \quad [2]$$

Using Equation [2], Allan variance was calculated for each wavenumber in the spectrum and a mean over the spectrum calculated to represent the average Allan variance for that time combination.

To characterise the rate of change of background suppression the 100 x 0.1 s, 50 x 0.2 s and 20 x 0.5 s time combinations were used. The total intensity across all wavenumbers was calculated for each time combination at each location of each sample and a mean taken for all the IDH-MUT locations and all the IDH-WT locations. Mean spectra were divided by the individual exposure time in order to normalise the intensities and were then fitted to a third-order polynomial curve. The first derivative was calculated for

[vi] The only spectral processing was truncation (350 cm^{-1} to 1850 cm^{-1}) and cosmic ray removal. Cosmic rays were identified and removed manually using the WiRE software "Zap" function which allows small spectral areas to be removed and replaces with a straight line. There were only 4 cosmic rays in this data set.

each curve fit over the first 3 seconds, and the mean of this taken to represent the initial rate of decrease in intensity. The negative of the first derivative was plotted to give a positive value for the rate of change.

3.4.2 Results

To illustrate the different time combinations, Figure 3.15 shows the raw spectra for one set of time combinations on an IDH-WT tumour sample.

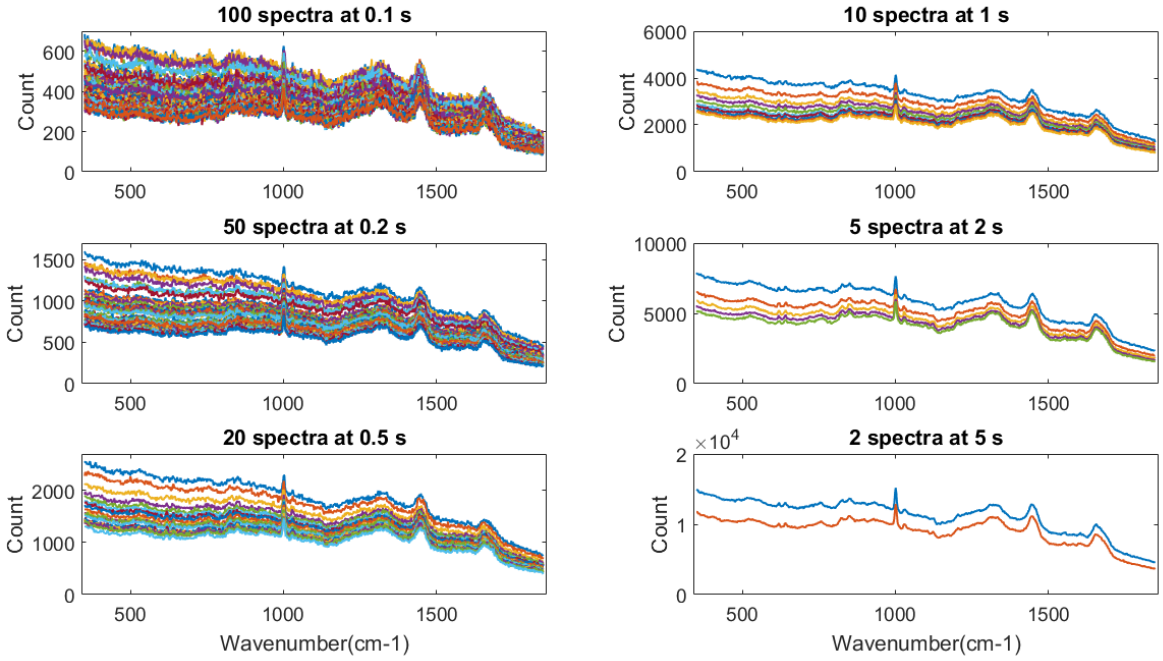


Figure 3.15 – Raw spectra for each exposure time combination for an IDH-WT tumour sample

With regards to characterisation of background suppression, Figure 3.16a illustrates the rate of change in intensity over 10 seconds for each of the 0.1 s, 0.2 s and 0.5 s time combinations. The IDH-MUT spectra have a greater starting intensity and a more rapid initial decrease in intensity compared to the IDH-WT spectra. In all time combinations, the change in intensity over the first 3 seconds was significantly different between the IDH-MUT spectra and IDH-WT spectra. At 0.1 s the rate of change in intensity (given in units of counts per second x 10^4) was 20.0 (95% confidence intervals (CI): 18.4 to 21.6) for IDH-MUT and 10.7 (95% CI: 9.7 to 11.8) for IDH-WT. At 0.2 s the rate of change was 18.6 (95% CI: 13.2 to 23.9) and 6.0 (95% CI: 2.4 to 9.6) for IDH-MUT and IDH-WT respectively. At 0.5 s the rate was 13.1 (95% CI: 10.0 to 16.2) for IDH-MUT and 4.6 (95% CI: 0.9 to 8.3) for IDH-WT (Figure 3.16b).

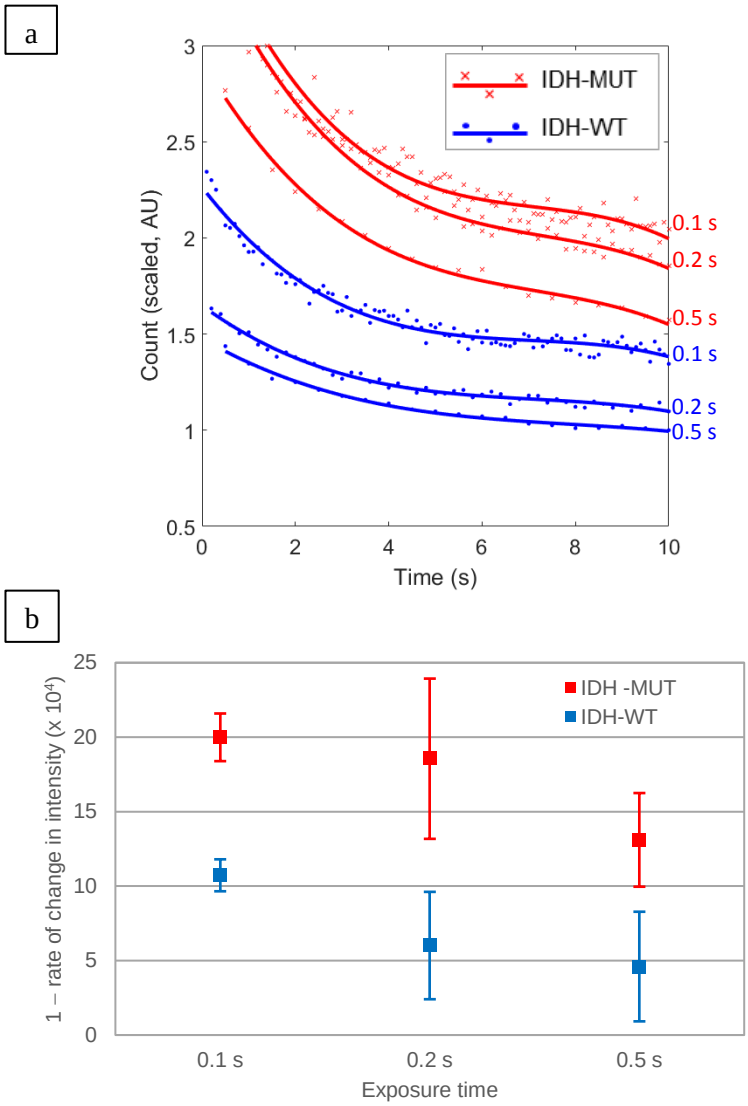


Figure 3.16 - Characterisation of the rate of intensity change. a) The normalised mean total intensity over time for IDH-MUT and IDH-WT samples for three different time combinations. b) Mean rate of change (mean negative first derivative) over the first 3 seconds for each time combination for IDH-MUT and IDH-WT samples (error bars= 95% confidence interval).

With regards to Allan variance analysis, Figure 3.17 demonstrates the summed Allan variance per wavenumber for each time combination for one sample. There is an initial decrease in Allan variance up to the 20 x 0.5 s time combination, after which the Allan variance increases. This is further illustrated in Figure 3.18, which shows the mean Allan variance for each time combination reaching a minimum at around 0.5 s to 1 s before increasing again as the exposure time increases. This relationship is reproducible in all locations and in both IDH-MUT and IDH-WT samples, although there is a hint that the minimum Allan variance occurs slightly later in the IDH-WT group (Figure 3.18b). This corresponds with the previous finding of a slower rate of change in intensity in the IDH-WT spectra.

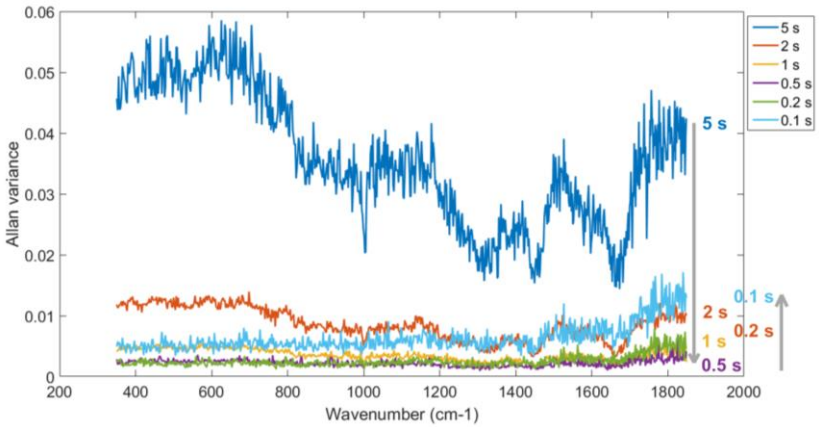


Figure 3.17 - Allan variance per wavenumber for each time combination for an IDH-WT tumour sample. Change in Allan variance is illustrated by the arrows.

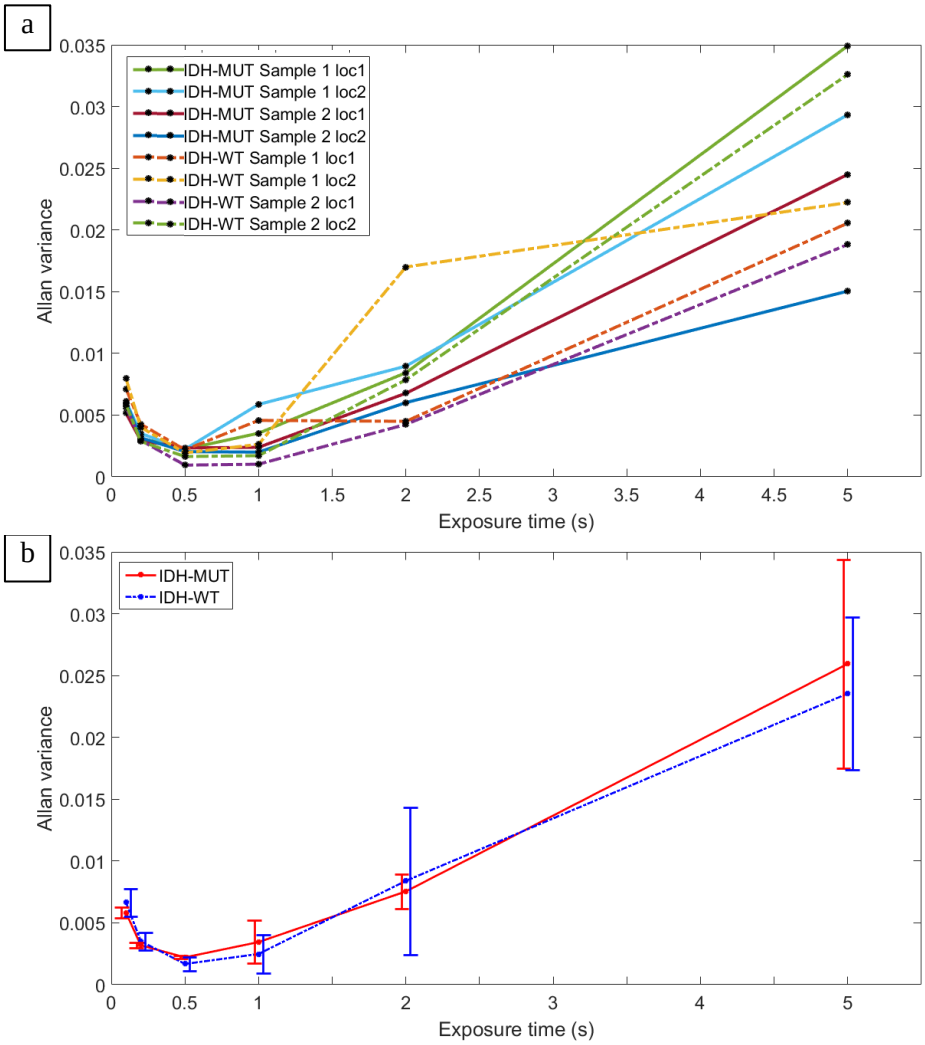


Figure 3.18 –Allan variance for IDH-MUT and IDH-WT samples. a) Mean Allan variance for each exposure time combination for all IDH-MUT and IDH-WT samples and locations. b) Mean for all IDH-MUT and all IDH-WT data in a) (error bars = standard deviation, x-axis offset of error bars for clarity)

Allan variance analysis was repeated using a point map area for each time combination rather than a single location (i.e. 0.1 s point spectra map repeated 100 times, 0.2 s point spectra map repeated 50 times… and so on) to confirm that there was no effect of the method of mapping. This yielded very similar results to those presented above, with the minimum Allan variance remaining at an exposure time of 0.5 s to 1 s.

3.4.3 Interpretation

These data support the hypothesis that a minimum in Allan variance is seen as the exposure time increases from very small values. This minimum should represent the optimum balance between providing sufficient laser exposure to record meaningful spectra with acceptable SNR and the possibility of laser induced changes in the spectra. It is important to note that this experiment cannot determine directly whether the increase in Allan variance at higher exposure times is in fact due to laser induced changes in the spectra, as changes in any other experimental parameters over a timescale of seconds would be expected to yield a similar result. However, the changes in intensity, increase in SNR and variations in IDH-MUT minus IDH-WT difference spectra seen with repeated laser exposure in this and the previous experiments would support this notion. Whatever the cause for the increase in variance, it is important to keep exposure time in the region of minimum Allan variance in order to limit the variability within the spectra of individual samples. In the present experimental setup, employing the RA800-series benchtop spectrometer, the optimum exposure time is found to be around 0.5 s to 1 s. To relate this back to the original preliminary 4 s and 6 s IDH data, undertaken on the inVia™ spectrometer, once differences in laser power and laser line area are factored in, one would predict an optimum exposure of between 1.2 s and 2.3 s[vii]. This comparison can only be considered an estimate, as the preliminary measurements used the StreamLine™ method of mapping which constantly changes the position of the laser line over the sample rather than sampling multiple individual point spectra. Despite this caveat, it is still clear that the 4 s exposure time in the preliminary data was already within the region

[vii] Calculated on the assumption that laser intensity (I) ∝ laser power (P) x exposure time (t) / laser line area (A)

P_{802} x t_{802} / A_{802} = P_{inVia} x t_{inVia} / A_{inVia}	therefore	t_{inVia} = P_{802} x t_{802} x A_{inVia} / A_{802} x P_{inVia}
P_{inVia} = 70 mW; A_{inVia} = 67 µm^2;	and	P_{802} = 245 mW; A_{802}= 100 µm^2.
t_{inVia} = 245 x t_{802} x 67 / 100 x 70 = 2.345 x t_{802}	for	t_{802} = 0.5 to 1.0, $\boldsymbol{t_{inVia}}$ **= 1.1725 to 2.345**

of possible laser induced spectral change, and this is likely to be even more of a problem in the case of the 6 s data.

The rate of intensity decrease over time was characterised as part of the Allan variance experiment and revealed a significant difference between the rate of background suppression in the IDH-MUT spectra compared to the IDH-WT spectra. This implies that the different tumour types are affected to varying degrees by increasing laser exposure. This provides further insight into the difference between the original 4 s exposure and 6 s exposure preliminary data, and supports the argument for keeping the exposure time at the point of minimum Allan variance.

3.5 Conclusion

In the work described in this chapter, several experimental variables have been explored in an attempt to explain the observed differences between data collected with a 4 s and 6 s exposure time in the preliminary genetic subclassification of gliomas experiment. The Raman spectra of FFPE samples have proven to be very stable, with no effect on the recorded spectra of laser exposure to nearby regions of the sample or of sample degradation over time. The spectra do, however, change as a result of direct laser exposure, with increasing exposure time not only causing a suppression of the background signal, but possibly also a change in the spectral characteristics. The different tumour types examined here were also found to respond to the laser exposure at different rates, which is also consistent with observations made in the preliminary genetic classification data.

From a practical point of view, several lessons have been learnt that will be important when employing the present experimental setup to perform the genetic subclassification of gliomas study.

- The exposure of samples to the excitation laser needs to be closely regulated. All further experiments therefore used simple point spectra maps with a step-size greater than the length of the laser line to ensure no overlap between illuminated regions within a mapping scan.
- For each sample preparation type (FFPE sections, cryosection or fresh tissue), change in sample preparation protocol, or change of spectrometer, the effect of laser power on the sample spectra should be evaluated via an Allan variance analysis.

- Once the optimum exposure time has been determined, this should be acquired over multiple accumulations. For example, if the optimum exposure time is 2 s then this should be made up of two times 1 s or four times 0.5 s accumulated exposures. Multiple accumulations will improve SNR, scaling as the square root of the total number of accumulations, but has been shown in this chapter to have an additional effect on increasing SNR because of the background suppression. An alternative method of dealing with the background suppression would be to expose the sample to the laser for a period of time prior to recording the Raman spectrum, allowing time for the background suppression, and the increase in SNR, to occur. This, however, would be limited to time intervals up to the optimum exposure time in order to limit any laser-induced spectral changes. By summing multiple accumulations over the optimum exposure time, there is essentially an averaging of the background from no background suppression, initially, to the maximum background suppression that occurs within the limits of the optimum exposure time. This represents a compromise between taking advantage of the SNR benefits of background suppression but without exceeding the optimum total exposure time determined by Allan variance.

The experiments in this chapter have to some extent characterised the phenomenon of background suppression. The presence of a broad background in Raman spectra of biological tissue is well recognised and a common reported artefact in biological Raman studies that often significantly masks the Raman signal[237,238,240]. There has been much focus on developing pre-processing background subtraction techniques after data collection to deal with the problem of background[243]. The cause of the background is commonly assumed to be tissue fluorescence[237,244] but this assumption is contrary to basic principles of autofluorescence. For fluorescence to occur an electron within a molecule needs to be excited to a higher lying electronic energy state and then return to the ground state by emitting energy in the form of a photon of light at a longer wavelength, observed as fluorescence (see Figure 1.3). In general, the lowest excited electronic state of molecules, from which fluorescence originates, have energies corresponding to ultraviolet and visible wavelengths (300-700 nm). Near-IR wavelengths, such as 785 nm, in general do not have enough energy to excite electrons to higher electronic states and therefore are less likely to result in fluorescence[164,238].

Indeed, this is the rationale for using near-IR wavelengths, such as 785 nm and 830 nm, when undertaking Raman measurements on biological tissue[164]. The excitation energy needs to be much higher to promote electrons to the higher lying energy levels from which fluorescence can occur. Biological fluorophores (compounds that fluoresce upon excitation) have been shown to emit little or no fluorescence at excitation wavelengths longer than ~600 nm[237,244,245].

The physical origin of the background seen in near-IR Raman spectroscopy is therefore not fully understood. There is very little published literature that addresses this, with the majority of relevant work coming from Hugh Byne's group. In one paper this group describe experiments in which Raman spectra were recorded using a 785 nm excitation laser, firstly for a fine powder of pure kappa-elastin protein and then having compressed the powder into a pellet[239]. A significant reduction in the background and increase in SNR was observed for the pellet relative to the powder. The authors suggest that the broad background cannot be due to fluorescence, as Kappa-elastin is transparent at 785 nm and the background and Raman SNR were dependent on the sample morphology and not chemical composition. Based on Raman spectra recorded for skin FFPE sections, they further showed that background suppression is possible, without chemical alteration in the spectra, with the use of a water immersion lens. They suggest that this further supports the theory that the background is due to a scattering process, since the immersion lens reduces the amount of stray scattered light due to the reduced refractive index difference at the tissue/water interface in comparison to the tissue/air interface.

In a further study, the group recorded a series of multiple 10 s exposures on FFPE skin sections, similar to the experiments reported in this chapter, and reported a similar decrease in background intensity and improvement in SNR[240]. They also reported no visual damage to the tissue and no change to the Raman signal. Once again, the benefit of an immersion lens was demonstrated, which immediately suppressed the background with no further background suppression on repeat exposures. The suppression of background seen with repeat laser exposure is assumed by many to be a "photobleaching" effect as the presumed fluorescence gets quenched. However, with the suggestion by Bryne *et al.* that the background is attributed to scattering of laser light (due to tissue morphology) rather than fluorescence (due to chemical composition), the background suppression is attributed to a photothermal "annealing" effect which, although not visible microscopically, results in local changes to the sample morphology and

therefore a reduction in scattering. They suggest that the immersion lens provides a heat sinking environment which reduces the photothermal effect and thus prevents decay in background.

The use of an immersion lens was trialled as part of the study presented here (Figure 3.19). Measurement of consecutive maps was repeated using a x100 dry and x100 water immersion lens on different sample locations and reproduced the findings of Byrne *et al.* described above, with immediate suppression of the background. However, in contrast to Byrne et al, no increase in SNR was seen. This may have been due to the nature of the immersion lens used. Given the lack of increase in SNR and the more intricate experimental setup, requiring the careful application of water droplets to the sample – not conducive to protocol translation into a clinical setting – the use of an immersion lens was not pursued in this study and thus is not reported fully.

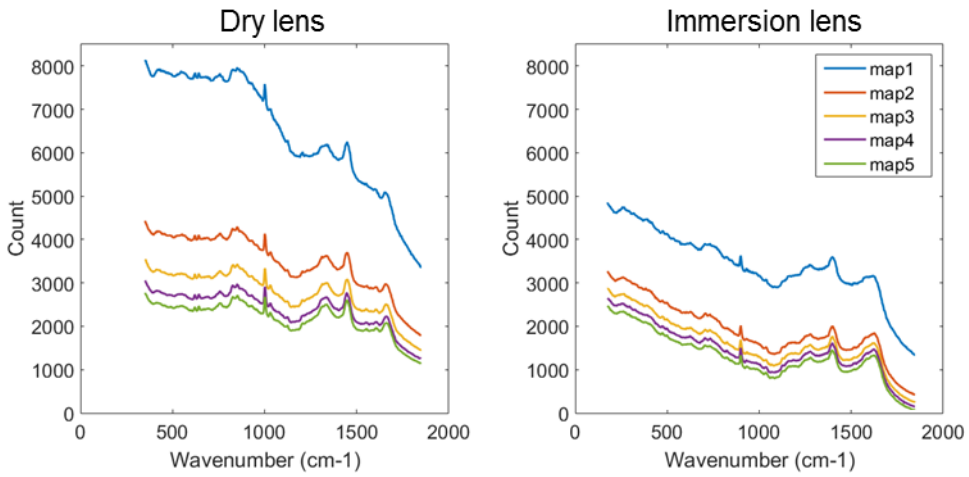


Figure 3.19 – Comparison of dry and immersion lens effect on background and background suppression with repeated exposures.

Further work is needed to fully understand the cause of the background and to explore the theory of photothermal "annealing" as a cause for background suppression. However, this is beyond the scope of this current study. To look for changes in tissue composition due to increased laser exposure, multiple homogeneous parallel tissue sections could be subjected to proteomic and metabolomics profiling after different durations of laser exposure in order to characterise any chemical changes in the sample. The combination of mass spectrometry to confirm entities measured with vibrational spectroscopy has been previously reported[246] and would be well suited to investigate these potential changes. The photothermal heating effect could be investigated by temperature-controlled experiments to see if there is an effect of

heating or cooling on the intensity suppression. Different substrates with varying thermal conductivity could also be investigated. Lastly, electron microscopy to view the tissue surface following differing extents of laser exposure would give a more detailed view of any possible physical change to the sample surface due to an annealing process.

Chapter 4. Rapid molecular genetic classification of gliomas using Raman spectroscopy

4.1 Introduction

As outlined in Chapter 1, the genetic classification of gliomas is critical for integrated neuropathological diagnosis and clinical decision-making. Moreover, there is increasing evidence that the impact of the extent of tumour resection also depends on the tumour genetic subtype (Section 1.4.3). Thus, information on the tumour genetic subtype at the time of surgery, providing real-time predictive and prognostic information, would allow the surgeon to tailor their surgical strategy to each patient's genetic profile. This chapter reports on a study using Raman spectroscopy to classify gliomas into their common genetic subtypes.

The main aim of this study was to evaluate the ability of Raman spectroscopy to identify the most common molecular genetic subtypes of diffuse glioma in the neurosurgical theatre setting. An important objective was to develop a simple protocol which would allow the surgeon or scrub staff to undertake Raman analysis of tissue in the operating theatre without the need for specialised technicians or laboratories. The application of Raman analysis in the neuropathological diagnostic workflow of glioma tissue, including analysis of cryosections, FFPE sections, and glioma cell lines (LN18) was also evaluated, with the key objectives being to use a range of preanalytical tissue preparations to validate the fresh tissue results, and to establish the utility of Raman spectroscopy in the neuropathology laboratory setting.

4.2 Methods

4.2.1 Study participants and tissue sample selection

Ethical approval for use of fresh, cryosectioned, and FFPE tissue was obtained through the Oxford Brain Bank generic Research Ethics Committee approval (15/SC/0639). All tissue used was deemed surplus to diagnostic requirements. Fresh tissue was obtained from all consenting patients between January 2017 and March 2018 as part of the routine intraoperative neuropathological evaluation. A new "fresh tissue for research" pathway was designed and implemented within the neurosurgery and neuropathology

departments to allow for safe fresh tissue analysis straight from the operating theatre without jeopardising the patients' routine diagnostic pathway. The fresh tissue for research standard operating procedure is shown in Appendix B. Cryosections and FFPE samples were acquired through application to the Oxford Brain Bank and were selected based on the availability of definitive neuropathological molecular and genetic diagnosis, the amount and quality of tissue available, and the appropriate patient consent.

4.2.2 Pathological diagnosis

All cases were reviewed by a Consultant Neuropathologist, and diagnoses were established using the EANO guidelines[38]. WHO integrated diagnosis was based on the IHC/genetic results and histological grade (Figure 4.1a), giving 5 possible diagnostic groups:

1. diffuse or anaplastic astrocytoma, IDH-mutant (WHO II or III);
2. glioblastoma, IDH-mutant (WHO IV);
3. oligodendroglioma or anaplastic oligodendroglioma, IDH-mutant and 1p/19q-codeleted (WHO II or III);
4. diffuse or anaplastic astrocytoma, IDH-wildtype (WHO II or III);
5. glioblastoma, IDH-wildtype (WHO IV).

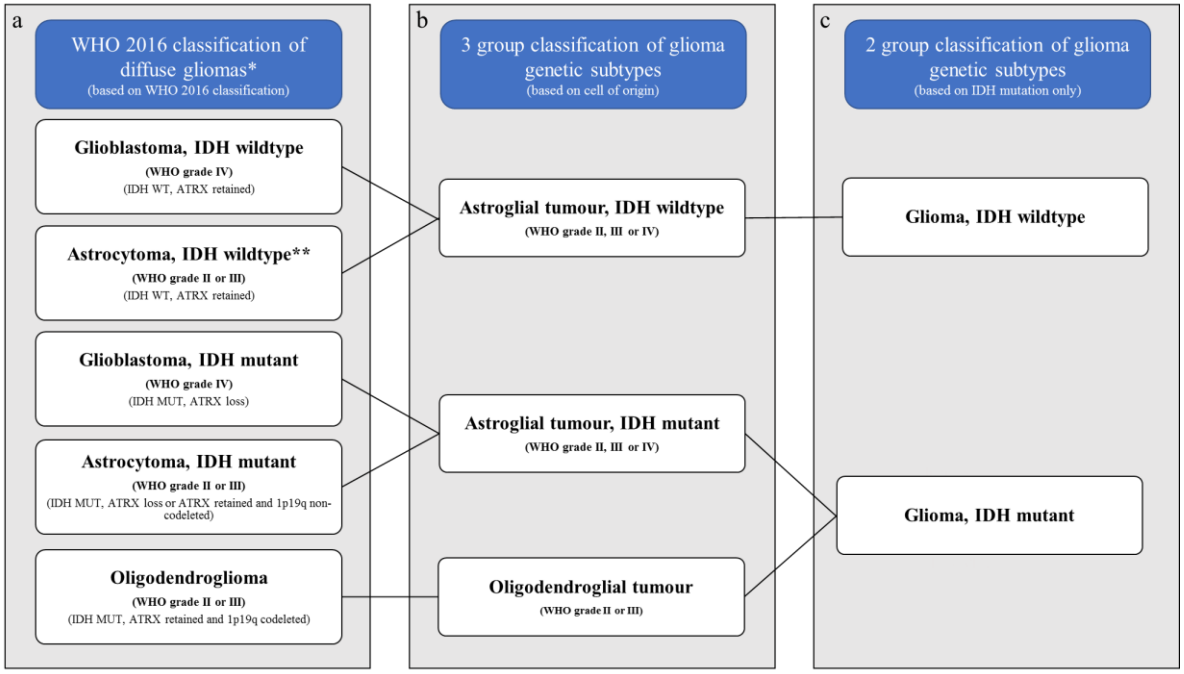


*Figure 4.1 – Pathological diagnoses and classification models. a) WHO 2016 classification of diffuse gliomas; b) genetic subtypes for three-group model; c) genetic subtypes for two-group model. (*not all subtypes are shown, only common groups; ** these are provisional categories in the classification: it is now recognised that most of these neoplasms are incipient or undersampled glioblastomas, IDH-wildtype*

IHC staining was carried out for the common IDH mutation (IDH1-R132H) and nuclear ATRX. FISH analysis for 1p/19q-codeletion was undertaken in cases that were IDH mutant and retained ATRX. As per EANO guidelines, IDH-wildtype and ATRX-retained cases in patients over 55 years old with characteristic glioblastoma histology were diagnosed as glioblastoma, IDH-wildtype, as the possibility of a rare IDH mutation being present is extremely low. Although not included in EANO guidelines, IHC for BCAT1 (branched-chain amino acid transaminase 1) was also carried out on the majority of cases. BCAT1 up-regulation has been strongly associated with IDH wildtype tumours[247]; in cases negative for IDH1-R132H IHC with diffuse overexpression of BCAT1 in tumour cells, the likelihood of a rare IDH mutation is close to zero[viii]. In all cases where BCAT1 is not up-regulated, or where there is histological or clinical ambiguity, the samples underwent targeted genetic sequencing for IDH1, IDH2, Histone H3F3A and BRAF V600E. Tumours that did not fit into one of the above categories (e.g. "Diffuse midline glioma" or "glioma, not otherwise specified (NOS)") were excluded from the study.

[viii] Local departmental observation of unpublished data (n=200)

4.2.3 Development of sample preparation protocols

Cryosection and FFPE section sample preparation protocols were developed based on the experimental optimisation experiments reported in Chapter 2. Much of this experimental development was also applicable to the fresh tissue samples. In addition, the fresh tissue preparation protocol was trialed using fresh pig brain and using the first three cases in the study as pilot cases[ix]. This was particularly important to test the mounting of fresh tissue on to a grooved stainless steel slide.

The aim was always to optimise the samples for Raman analysis whilst limiting the need for complex laboratory procedures. The latter is particularly important in the case of fresh tissue samples, for which the eventual aim is to develop a point-of-care protocol that can be employed in neurosurgical theatres. The LN18 cell line batches, once harvested, were processed as either cryosections or FFPE sections, with protocols to match the respective tissue sample preparations.

The finalised sample preparation protocols for each tissue type are outlined below.

Fresh tissue

Biopsies were taken by the neurosurgeon from an area of tumour bulk defined both visually and using the intraoperative neuronavigation system (Brainlab, Germany). The biopsy was immediately transferred to the neuropathologist, who cut the tissue biopsy into two pieces: one for intraoperative smear evaluation; and the adjacent piece for Raman analysis. The portion for Raman analysis measured approximately 3 mm^3. The sample was immediately transferred to a grooved stainless steel slide and compressed into the groove using a surgical scalpel (Figure 4.2). This created a relatively flat surface onto which the Raman laser could be focused, as well as compacting the tissue to limit the effect of tissue drying. Stainless steel was used due to the high SNR advantage of this substrate, as reported in Chapter 2. In addition, stainless steel is very strong, durable and can be cleaned and sterilised easily which makes is suitable for the sterile neurosurgical environment. The adjacent biopsy underwent intraoperative smear preparation with haematoxylin and eosin (H&E) or methylene blue staining and was reviewed by the Consultant Neuropathologist to confirm the presence of tumour. If there was evidence of necrosis or

[ix] Spectral data from the pilot cases were evaluated in terms of quality prior to being included in the main study, and all three datasets were found to be of sufficient quality to be included in the main study.

normal brain on these preparations, the Raman spectra from that sample were excluded from data analysis and a further fresh tissue sample was requested from the neurosurgeon if possible.

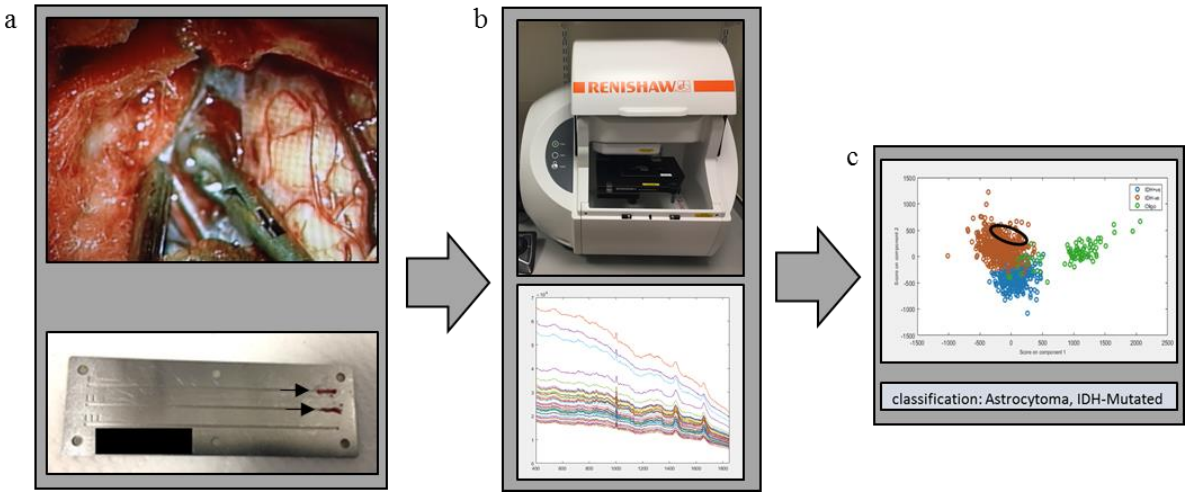


Figure 4.2 - Fresh tissue workflow. a) Intraoperative fresh tissue biopsy pressed into grooved stainless steel slide (arrows). b) Slide placed in Raman spectrometer and multiple spectra acquired. c) Spectra classified into genetic subtype using PC-LDA model.

Cryosection

Two parallel 10 µm sections were cut from selected samples using a cryotome. One section was mounted on a glass slide and underwent standard H&E staining. The second section was mounted on an APTES coated (Appendix A) stainless steel slide and returned to the –80 °C freezer. When Raman measurements were to be undertaken, these sections were removed from the freezer and allowed to thaw for 15 minutes. H&E sections were reviewed by a Consultant Neuropathologist, and areas of tumour identified. These areas were then identified on the unstained parallel section on the stainless steel slide, and Raman maps were recorded in these locations. If the H&E sample did not demonstrate evidence of tumour (usually indicating normal brain, necrosis or haemorrhage) then the Raman sample was excluded from the study.

FFPE

Two parallel 6 µm sections were cut from selected cases using a microtome. One section was mounted on a glass slide and underwent standard H&E staining. The second section was mounted on an APTES coated stainless steel slide (Appendix A). These were then placed in an oven at 40 °C for 30 minutes to increase sample adherence to the slide. Following this, samples were deparaffinised using Histo-Clear (National Diagnostics, USA) incubations for 3 x 5 minutes. Samples were then incubated with 100% isopropyl alcohol (IPA) for 2 x 10 dips, rehydrated (90% IPA for 10 dips, 70% IPA for 10 dips, distilled

water for 10 dips) and allowed to air dry before being stored in a slide box until they were used for Raman analysis. Similar to the cryosections, H&E sections were reviewed by a Consultant Neuropathologist and areas of tumour identified and located on the unstained stainless steel section. Samples not demonstrating evidence of tumour were excluded from the study.

LN18 cell line

Mutagenised and cloned R132H human IDH1 was cloned into a transfer plasmid of lentiviral vectors which were then transduced into the LN18 parental cells (ATCC, UK). The LN18 transduced cells overexpress IDH-R132H transgene compared to the IDH wildtype parental cells. Mutated and parental cell lines were then grown in parallel for two weeks in Dulbecco's modified Eagle media prior to harvesting. Cells were detached using trypsin to keep cells intact and then washed with phosphate buffer solution (PBS)[x]. The cells were divided into two batches: one for cryosections and one for FFPE sections, with the aim of establishing protocols similar to those used for tissue sample preparation. Cells destined for FFPE sections were placed directly in 10% formalin for 12 hours after harvesting. They were then centrifuged (2000 rpm for 5 minutes) to form a pellet, which was transferred to an embedding capsule and embedded in paraffin overnight. Three parallel sections of 6 μm thickness were cut from the cell paraffin blocks. One section was placed on a stainless steel slide for Raman analysis and underwent the same deparaffination and preparation as the FFPE tissue sections. Of the remaining two sections, one underwent H&E staining and one IDH1-R132H immunohistochemistry to confirm the presence or absence of the mutation in the cell cultures. Cells destined for cryosection preparation were suspended in phosphate buffer solution after harvest and centrifuged (1300 rpm for 5 minutes) to form a pellet. The supernatant was removed and an embedding medium (OCT (CellPath Ltd, UK)) added to the Eppendorf tube and allowed to settle at the bottom over the cell pellet. The pellet was snap frozen using isoprolene cooled with liquid nitrogen for 1 to 2 minutes. Once frozen, the Eppendorf tube was cut with a hacksaw just above the frozen pellet, which was then removed and mounted on cork with the embedding medium. The mounted pellet was snap frozen again in isoprolene and then stored in a –80 °C freezer. Three

[x] All cell culture preparation up to this point in the protocol was carried out by Joan Gannon and John Walsby-Tickle from Prof. Chris Schofield's and Prof. James McCullagh's groups, Department of Chemistry, University of Oxford.

parallel sections of thickness 10 μm were cut from the cell cryoblocks. One section was placed on a stainless steel slide for Raman analysis as per the cryosection tissue samples. The other two sections were mounted on standard glass slides and underwent H&E staining or IDH1-R132H immunohistochemistry.

Two LN18 IDH-mutant and two IDH-wildtype cell lines were cultured and processed separately at two different time points. Predictive Raman models were built for each cell line and also for the combined spectra from both cell lines.

4.2.4 Development of Raman spectroscopy acquisition protocols

Spectrometer specifications

The Renishaw RA800-series benchtop spectrometer was installed in a purposely redesigned room in the Neuropathology Department next to the operating theatres, allowing for rapid Raman analysis of fresh tissue samples. The spectrometer is equipped with an ×50 objective and a 785 nm excitation laser with 180 mW power at the objective. The spectrometer has automated auto-alignment and self-calibration sequences, based on built-in neon and silicon references. These were employed before all Raman data collection in this work. Data acquisition parameters were controlled using WiRE v4.3 software (Renishaw, UK). The spectrometer auto-calibration and simple user interface lends itself to clinical application and thus clinical translation of this technique.

Justification for protocol used

The methods and conclusions drawn from the experiments using FFPE samples reported in Chapter 3 were used to establish the optimum acquisition parameters to achieve the highest possible signal-to-noise ratio without CCD saturation, tissue damage, or change in spectral characteristics due to laser exposure. With regard to the effect of laser exposure on surrounding tissue (see section 3.2), there was no effect observed on spectra recorded at sample positions 500 μm apart in either cryosections or fresh tissue[xi], as observed with FFPE sections. To prevent overlap of laser exposure, a spacing of at least 500 μm was maintained between each area mapped on a sample. Additionally, within each map, the distance between

[xi] For fresh tissue the effect of laser exposure on surrounding tissue was investigated using a grid of single spectra, rather than a grid of maps (see figure 3.4), due to the tendency of fresh tissue to saturate the CCD and the sample to burn with repeated laser exposure.

each point spectrum (known as the step size) was always greater than the dimensions of the laser line. The RA800 excitation laser has dimensions of 80 μm by 2 μm along the *x* and *y* axes, respectively; thus the step size in all maps was set to 100 μm by 50 μm along the two axes.

For investigation of sample degradation over time, Chapter 3 reports FFPE section spectra to be stable for at least 5 weeks. With fresh tissue and cryosections this time period has no practical relevance, and degradation was therefore investigated for up to 6 hours. Fresh tissue samples were stored in Ringer's lactate solution[xii] and kept in a 4 °C refrigerator. Cryosections were stored in a slide box once thawed and kept at 4 °C. Samples were removed for Raman analysis every two hours from 0 to 6 hours. There was no degradation observed in either fresh tissue or cryosection spectra over the 6 hour time period, and no effect of the Ringer's lactate solution on the fresh tissue Raman spectra[xii].

Repeated laser exposure was shown in Chapter 3 to suppress the spectral background and increase SNR up to a point. To this end all total exposure times were split into multiple accumulations. It was also shown that repeated laser exposure has a possible effect on spectral characteristics. The exposure time required to balance these two factors was estimated using Allan variance measurements (see section 3.4). It was found that cryosection tissue could be exposed to up to 10 to 15 s of 100% laser power without CCD saturation, considerably higher than the maximum of 5 to 10 s found for FFPE tissue. Allan variance measurements were therefore performed over a total of 15 s exposure time compared to the 10 s used in the FFPE experiments. The minimum mean total Allan variance was reached at around 7 to 9 s exposure time (see figure 3.18 to compare with FFPE). For fresh tissue samples the maximum laser exposure before CCD saturation was much lower than cryosections or FFPE sections and varied widely from 2 to 5 s. Fresh tissue was also prone to thermal damage, and resultant CCD saturation, with visual evidence of sample burning as shown in Figure 4.3. This thermal damage occurred at longer exposure times or, more commonly, when the biopsy was particularly bloodstained, most probably due to optical absorption by haemoglobin resulting in significant sample heating. Four out of the 62 fresh tissue samples used in the main genetic classification study had problems with thermal damage and tissue

[xii] Ringer's lactate solution is routinely used and readily available in neurosurgical theatres. Many of the fresh tissue samples came into contact with Ringer's lactate anyway as it is often placed in the biopsy specimen pot to help remove the biopsy from the biopsy forceps. 1 L Ringer's lactate (Osmolarity 273 mOsm L^{-1}) contains: 130 mmol L^{-1} sodium, 109 mmol L^{-1} chloride, 28 mmol L^{-1} lactate, 4 mmol L^{-1} potassium, 1.5 mmol L^{-1} calcium.

burning. In every case, however, it was possible to acquire good quality spectra from non-bloodstained regions of the biopsy. Allan variance measurements could therefore only be carried out for a total of 2 s exposure time and showed no evidence of spectral changes over the 2 s exposure. Despite restrictions in the amount of laser power that could be used, good quality spectra with excellent signal to noise ratios where possible with 0.5 to 1 s exposures. This was likely due to the thickness of the samples compared to sections cut with a microtome.

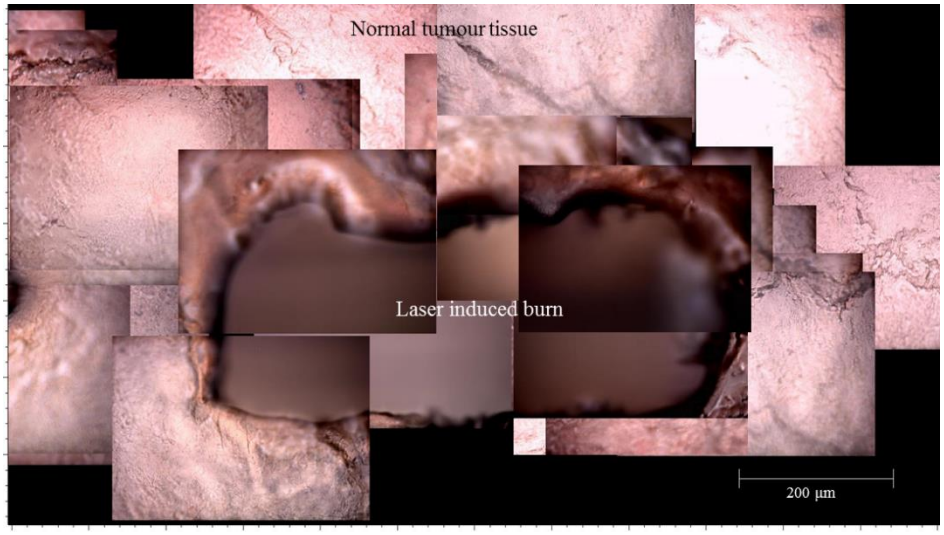


Figure 4.3 - Laser induced burn on fresh tissue sample. Montage image created at x50 zoom.

A further challenge when working with fresh tissue was maintaining a well-focused image throughout a mapping sequence. The method of pressing the tissue into a grooved slide created a flat surface, which made coarse focusing easily achievable without the risk of contaminating or damaging the objective lens. Fine focusing was more difficult, as the spectrometer did not have an autofocus function, and was done manually by the user. Areas on the sample of similar focus were easily found in all cases, but frequently the sample went out of focus as the area was being mapped by the laser. It was observed that to refocus on the sample the microscope stage always needed to be moved upwards (increasing the *z* coordinate of stage position and reducing the distance between the sample and the microscope objective), and it was therefore hypothesised that the change in focus was due to drying of the sample as it was exposed to the laser, causing it to retract slightly into the slide groove and thus move away from the objective. This explanation was further supported by the observation that the defocusing effect became less apparent over the few minutes that the sample was being analysed, presumably as the rate of sample drying

decreased. To estimate the effect of sample drying, and thus defocusing, on spectral SNR, a series of experiments were carried out. A series of spectra were recorded at the same location on a freshly prepared tissue sample at 0.25 s intervals over 0 to 3 s with a 0.1 s exposure time. This was repeated at 3 locations on the sample, and a mean value of the percent change in SNR[xiii], relative to the initial spectrum, was taken for each time point. This was also repeated once the data collection for the main study presented here had been collected, i.e. after the sample had dried out. Figure 4.4a shows the relative change in SNR over the 3 seconds for measurements taken as soon as the sample was placed in the spectrometer ("Early") and for measurements taken at the end of main data collection ("Late"). In support of the drying hypothesis, there is a greater loss in relative SNR in the Early compared to the Late measurements. Overall, the Early measurement SNR loss is greater than 25% by 3 s. In the Late measurements there is a slight increase in SNR initially which is consistent with the background suppression, and resultant increase in SNR, reported in Chapter 3, although the magnitude of this variation is small and may simply represent noise. These data all support the drying hypothesis. To refocus the sample at the end of the Early measurements an average increase in the microscope stage z-axis of 17 μm was required, compared to 4 μm in the Late measurements.

To establish how the SNR changes as a function of focal depth, single spectra were acquired with the microscope stage *z*-axis at varying distances from the focal point (0.5 s exposure time, new location on sample for each acquisition). Negative *z* values moved the focal point above the sample, positive z values moved the focal point into the sample. Figure 4.4b shows the relative percent change in SNR as a function of *z*. As the focal point moves above the sample (negative *z* values, microscope stage moved down) the relative SNR drops off rapidly. This is to be expected because less Raman scatter will be produced, as less laser power is focused on the sample, and less Raman scatter is being collected by the objective. As the focal point moves into the sample (positive *z* values, microscope stage moved up) the drop in SNR is more gradual with a possible initial increase in the SNR. Again, this is to be expected because initially the laser and resulting Raman scatter will pass through the tissue to/from the focal point with the addition of extra signal from excitation of tissue within the cone of light above the focal point (this is also the

[xiii] SNR was calculated using phenylalanine peak at 1003 cm^{-1} (signal range: 980-1020 cm^{-1}, noise range 930-980 cm^{-1}).

explanation for the high SNR achieved with relatively low exposure times in fresh tissue compared to cryosections for FFPE sections). Once the tissue thickness results in too much absorption of the laser and Raman scatter the SNR decreases. These and the previous experiment's data demonstrate that there is a narrow window of focus in which SNR is maintained, and positioning the focal point 2 or 3 µm beyond the surface, into the sample, would be advantageous. For future clinical applications, a point-of-care device would require an autofocus function (recently available in this series of spectrometers (LiveTrack™, Renishaw, UK)) and, given the narrow focal window, the automated system would need to refocus on the sample after every acquisition. Refocusing after every spectrum was not a practical solution in the main study reported here, and with the current setup as the aim was to acquire as many

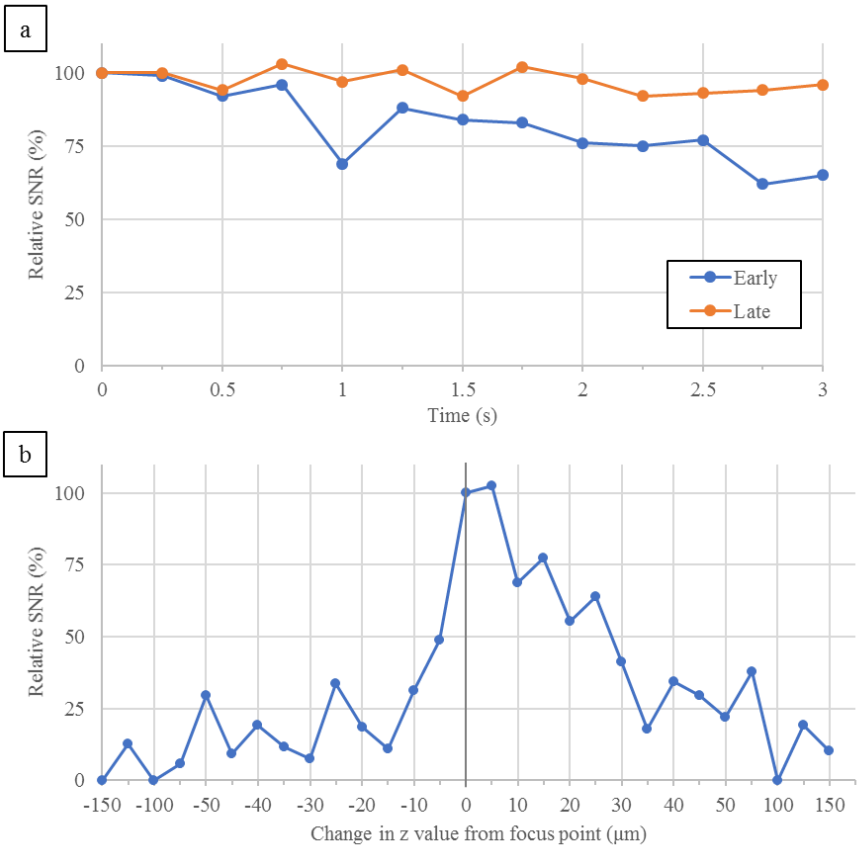


Figure 4.4 – Effect of focus on SNR. a) Change in relative SNR with repeated 0.1 s exposures every 0.25 s. b) Change in SNR versus change in focal point (microscope stage z-axis).

spectra from the sample as possible. Instead, small area point maps, acquiring approximately 20 to 30 spectra, were measured and the focal point was positioned up to 5 µm beyond the surface focus with the focus visually checked at the end of the point map.

Final Raman spectroscopy acquisition protocols

For cryosections and FFPE samples, two or three (depending on the size of sample available) 0.5 mm^2 areas were chosen for Raman analysis corresponding to the area identified by the Neuropathologist on the parallel H&E section as containing tumour. A square grid of 66 single spectra (a point map) was acquired over each chosen area. Figure 4.5 shows part of the WiRE software point map setup window with an x-axis step size of 100 µm and z-axis step size of 50 µm giving 66 spectra distributed in a 6 by 11 grid within the 0.5 mm^2 area.

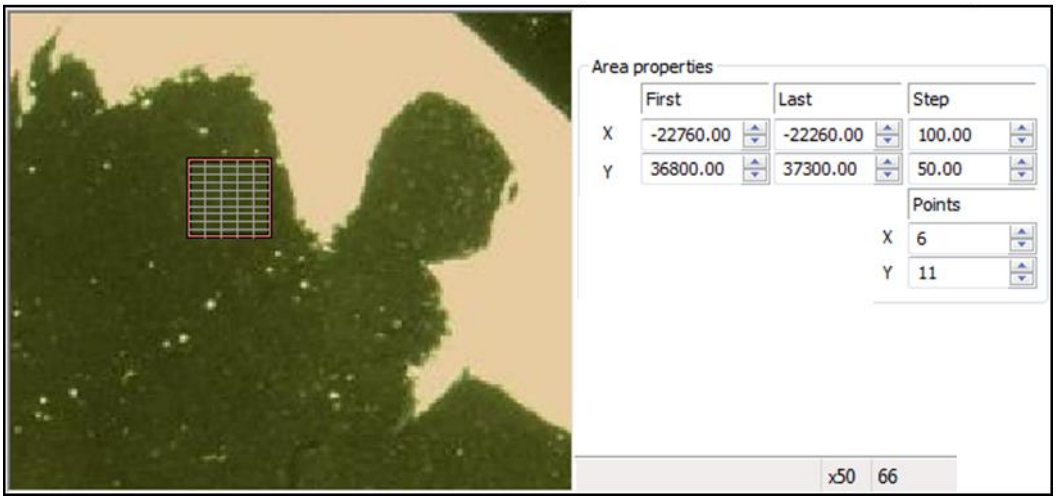


Figure 4.5 - Cryosection and FFPE point map setup. Point map setup window in WiRE software. X and Y coordinates, step size (µm) and the resulting number of points in X and Y (and total, 66) are shown. Image shows sample at x50 magnification and location of point map. (Other menu options have been masked out for clarity)

Acquisitions consisting of two accumulations of 5 s for cryosection and two accumulations of 1 s for FFPE sections. For the LN18 cells, five 500 μm^2 areas of high cell density were chosen, using the same acquisition protocol as for the respective tissue cryosections and FFPE sections. For fresh tissue samples, the size of area analysed varied, to include between 20 to 30 spectra, depending on the size of an area that remained in focus at a single focus point. The same x- and z-axis step sizes were used as in the other tissue types. The aim was to obtain at least 100 spectra per fresh tissue sample. The acquisition protocol for fresh tissue was two accumulations of either 0.5 s or 1 s depending on the amount of background fluorescence and resulting CCD saturation for that sample. The total time to acquire data from each area was taken from the spectrometer log. The time from surgical biopsy to completion of spectral collection was also recorded.

4.2.5 Preprocessing of spectroscopic data

All spectral and data analysis was carried out using MATLAB R2015b (MathsWorks, USA). The aim of preprocessing is to remove components of the spectra that are not related to Raman scatter produced by the laser interaction with the substance being analysed. A balance must be struck when using preprocessing and modelling techniques to allow for filtering of unwanted components but not overfitting of the data, which would make the models less likely to maintain performance when test or validation datasets are used. To this end, in this study the effect that a preprocessing step had on the subsequent model performance was independently tested. For example, the effect on model performance of baseline correction was evaluated with polynomial orders ranging from 1 to 12. In the case of filtering methods, such as SNR thresholds, in no model was more than a total of 30% of the spectra removed by filtering.

The following preprocessing steps were carried out on the data:

- Spectra were truncated to include only the region between 400 and 1850 cm^{-1} wavenumbers. The range between 1560 to 1578 cm^{-1} was also excluded, as this overlaps with the Raman band of molecular oxygen.
- A SNR filter was applied to all spectra. In this study the SNR was calculated using the phenylalanine peak at 1003 cm^{-1}, with the signal region set to span 980-1020 cm^{-1} and the noise region spanning 930-980 cm^{-1}. The root mean squared (RMS) value of the noise region was used to estimate the noise. For the signal, the linear baseline across the signal region limits was calculated and used to baseline-correct the signal region; the resulting peak value represented the signal. The SNR was calculated from the ratio between the baseline-corrected signal and the RMS noise. A threshold SNR value was determined for each tissue type, and any spectrum with SNR below this threshold was excluded from the dataset. SNR thresholds used were 4.5 for fresh tissue, 6.0 for cryosections, and 8.0 for FFPE sections.
- A standard deviation threshold was found to be the most effective way to remove cosmic ray artefacts in these data[xiv]. The mean and standard deviation (STD) at each wavenumber for each

[xiv] Median filtering was also used as a method of noise and cosmic ray reduction. This involves calculating the median spectrum from a specified grid of surrounding individual spectra. To use median filtering the spectra at the edges of each map have to be removed (as they have no surrounding spectra to calculate the median). Given the small size of the maps used in this study too many spectra were being excluded using this technique.

spectral map was calculated. A STD factor was determined which set a STD range equal to the mean + STD*STD factor. Any spectrum with a spectral point outside the defined range was excluded from the dataset. The lower the STD factor the more spectra were rejected. Spectra showing CCD saturation were filtered using both the SNR threshold and the STD threshold. STD threshold values used were 6 for fresh tissue, 4.5 for cryosections, and 3.5 for FFPE sections.

- Spectral baseline correction and normalisation was carried out as part of Extended Multiplicative Scatter Correction (EMSC). EMSC is a model-based preprocessing method regularly used in vibrational spectroscopy to normalise and background correct data in a single operation[248]. The EMSC model used in this study assumed each spectrum to be made up of 1) a background signal; 2) the spectral contribution of the objective lens; 3) the mean spectrum of all the dataset and 4) the unique aspects of the spectrum of interest. The data were normalised to the mean of the entire dataset, while the background was modelled by a measured objective spectrum combined with a polynomial baseline (a third-order polynomial was used throughout).
- For model building, the central region of each preprocessed spectrum, excluding 50 cm^{-1} from each end, was used in order to minimise the influence of any edge artefacts from the polynomial background correction.

4.2.6 Classification model building

Principal Component Analysis fed Linear Discriminant Analysis (PC-LDA) was used to build all pathology classification models in this study. PCA is an unsupervised learning method that uses an orthogonal transformation to convert a number of possibly correlated multidimensional variables into a smaller number of uncorrelated variables called principal components (PC)[249]. Each successive PC gives the direction in the data that represents the maximum variance remaining in the dataset. PCA is used here to reduce the data dimensionality. Analysis of variance (ANOVA) was used to select the PCs that describe a statistically significant difference in the data for distinguishing between the different pathologies and these PCs were used in the LDA model. LDA is also a linear transformation technique,

but differs from PCA as it is a supervised learning method that computes the directions that will represent the axes that maximises the separation between defined multiple classes.

Model performance was assessed using leave-one-patient-out cross-validation to generate model sensitivity and specificity. Multiclass receiver operating characteristic (ROC) curves were generated and the area under ROC curve (AUROC) calculated as a measure of accuracy.

For each tissue preparation type, two separate models were built. The first was a three-group model which stratified the gliomas into the three main genetic subtypes regardless of histological grade (Figure 4.1b). The three-group model comprised of the following categories:

1. astrocytomas, IDH-mutant (astrocytoma (WHO II or III) and glioblastoma (WHO IV));
2. astrocytomas, IDH-wildtype (astrocytoma (WHO II or III) and glioblastoma (WHO IV));
3. oligodendrogliomas, IDH-mutant and 1p/19q-codeleted (WHO II or III).

The second model stratified based on the IDH mutation alone in order to compare performance with the neuropathology laboratory immunohistochemistry IDH assay (Figure 4.1c). This two-group model comprised of:

1) IDH-mutant (including all IDH-mutant astrocytomas, glioblastomas and oligodendrogliomas);
2) IDH-wildtype (including all IDH-wildtype astrocytomas and glioblastomas).

All astrocytomas, IDH-wildtype were glioblastomas (WHO IV) – there were no grade II or III astrocytomas, IDH-wildtype tumours; these are generally accepted to be undersampled glioblastomas[36].

4.2.7 Pathology spectra representation and peak assignment.

The mean spectrum of each genetic subtype was calculated after preprocessing (including EMSC). These were plotted for the three-group and the two-group model for each tissue preparation type. The differences between these spectra are often very small and difficult to visualise, and thus the second-derivative transformation of the mean spectrum was derived and the negative of this (-2ndD) was plotted in order to better resolve the Raman peaks[250] [xv]. Using these plots, tentative molecular assignment of the identified Raman peaks was carried out based on published literature[176,182,231-233]. For each wavenumber

[xv] It is important to note that the -2ndD plot only represents the location of the Raman peaks and the plot does not indicate the difference between the mean spectra of different pathological types. The latter can be evaluated using difference spectra.

at which a Raman peak was identified, the normalised intensities of all spectra of a given genetic subtype were analysed and compared with the intensities of the other genetic subtype(s) in the model. In both the two-group and three-group models, IDH-mutant intensities were compared with IDH-wildtype intensities. Additionally, in the three-group model, oligodendroglioma intensities were separately compared with astrocytomas, IDH-mutant and astrocytomas, IDH-wildtype intensities. Statistical differences between these intensities were tested using the Wilcoxon rank sum test, and the magnitude of the resulting *p* value was used as an indicator of the magnitude of the difference between the intensities.

4.3 Results

A total of 62 fresh tissue samples were included in the study. No cases were excluded on the grounds of the absence of tumour on the adjacent H&E analysis of the smeared section. 97 cryosections and 132 FFPE sections were cut for analysis. Of these, 18 cryosections and 12 FFPE sections were excluded after review of the parallel H&E section showed no evidence of tumour. A total of 79 cryosections and 120 FFPE sections were used for analysis. Table 4.1 shows the number of samples analysed in each genetic subgroup for each sample preparation method.

<table>
<tr><th rowspan="3">Genetic subtypes
(WHO grade)</th><th colspan="6">Tissue preparation</th></tr>
<tr><th colspan="2">Fresh tissue</th><th colspan="2">Cryosections</th><th colspan="2">FFPE sections</th></tr>
<tr><th>3-grp model</th><th>2-grp model</th><th>3-grp model</th><th>2-grp model</th><th>3-grp model</th><th>2-grp model</th></tr>
<tr><td>Astrocytoma, IDH-wildtype
(All WHO IV (GB))</td><td>36</td><td>36</td><td>19</td><td>19</td><td>41</td><td>41</td></tr>
<tr><td>Astrocytoma, IDH-mutant
(WHO II / WHO III/WHO IV (GB))</td><td>21
(4 / 7 / 10)</td><td rowspan="2">26</td><td>41
(12 / 10 / 19)</td><td rowspan="2">60</td><td>51
(8 / 12 / 31)</td><td rowspan="2">79</td></tr>
<tr><td>Oligodendroglioma
(WHO II / WHO III)</td><td>5
(3 / 2)</td><td>19
(8 / 11)</td><td>28
(12 / 16)</td></tr>
<tr><td>Total</td><td colspan="2">62</td><td colspan="2">79</td><td colspan="2">120</td></tr>
</table>

Table 4.1 - Number of samples. Number of each genetic subtype included for fresh tissue, cryosections and FFPE sections, subdivided into 3-group model and 2-group model (GB= glioblastoma; grp= group).

9,799 spectra were collected in the fresh tissue group, with a mean of 161 spectra per sample. The mean time for data collection was 6 minutes and 6 seconds per sample, and the whole process from

intraoperative biopsy to completion of Raman spectroscopy was under 15 minutes in every case. In the cryosection group, 15,576 spectra were collected from 236 different 0.5 mm^2 regions on the sections (66 spectra per region). Each region took approximately 13 minutes to complete. In the FFPE sections, 23,760 spectra were collected from 360 0.5 mm^2 regions (66 spectra per region). Each region took approximately 3 minutes 30 seconds to collect.

4.3.1 Classification model performance

Table 4.2 summarises the sensitivities and specificities for three-group and two-group prediction of glioma genetic subtypes using PCA-LDA classification modelling of the Raman spectra with leave-one-patient-out cross validation. LDA component plots and ROC curves for these models with fresh tissue, cryosection and FFPE sections are illustrated in Figure 4.6, Figure 4.7 and Figure 4.8, respectively.

With fresh tissue samples, the models were able to classify tumours into either IDH-mutant or IDH-

Model	**Genetic subtypes**	**Fresh tissue**		**Cryosections**		**FFPE sections**	
		sens	spec	sens	spec	sens	spec
3-group	**Astrocytoma, IDH-wildtype**	0.94	0.90	0.78	0.85	0.81	0.84
	Astrocytoma, IDH-mutant	0.91	0.95	0.79	0.89	0.72	0.87
	Oligodendroglioma	0.79	1.00	0.74	0.90	0.79	0.93
2-group	**IDH-wildtype**	0.95	0.91	0.88	0.83	0.77	0.71
	IDH-mutant	0.91	0.95	0.83	0.88	0.71	0.77

Table 4.2 - Glioma classification model performance. Sensitivity (sens) and specificity (spec) for predicting genetic subtypes using PC-LDA 3-group and 2-group models for fresh tissue samples, cryosections and FFPE sections.

wildtype (2-group model) with 0.91 to 0.95 sensitivity and specificity and 0.98 accuracy (AUROC, Figure 4.6d). Fresh tissue classification into the three genetic subtypes (3-group model) was performed with 0.79 to 0.94 sensitivity, 0.90 to 1.00 specificity and an accuracy of between 0.92 to 0.98 (Figure 4.6b). Cryosection and FFPE section models classified with slightly less sensitivity and specificity than the fresh tissue models, as shown in Table 4.2. The overall accuracy of these models remained high, with cryosection three-group classification accuracy of 0.89 to 0.92 (Figure 4.7b), cryosection two-group classification accuracy of 0.93 (Figure 4.7d), FFPE section three-group classification accuracy of 0.87 to 0.93 (Figure 4.8b), and FFPE section two-group classification accuracy of 0.82 (Figure 4.8d). LDA

plots for the three-group and two-group models show good separation of the genetic subtypes using fresh tissue (Figure 4.6a and c), cryosections (Figure 4.7a and c), and FFPE sections (Figure 4.8a and c).

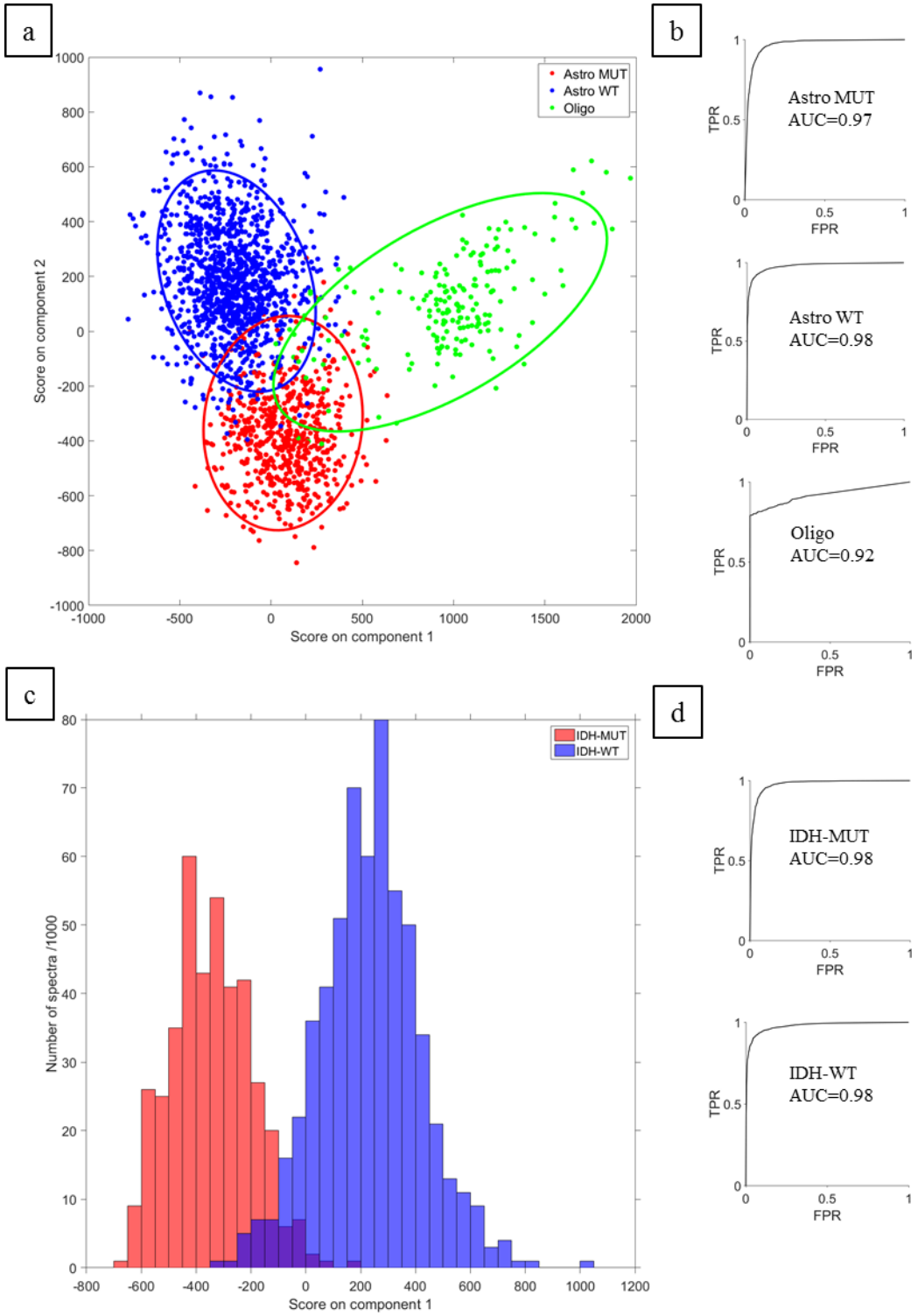


Figure 4.6 – Fresh tissue model performance. Fresh tissue LDA score plots and ROC curves for 3-group model (a, b) and 2-group model (c, d). (TPR= true positive rate; FPR= false positive rate; Astro-MUT = astroglial tumour, IDH mutant; Astro-WT = astroglial tumour, IDH wildtype; Oligo = Oligodendroglioma; IDH-MUT = IDH mutant tumour; IDH-WT= IDH wildtype tumour).

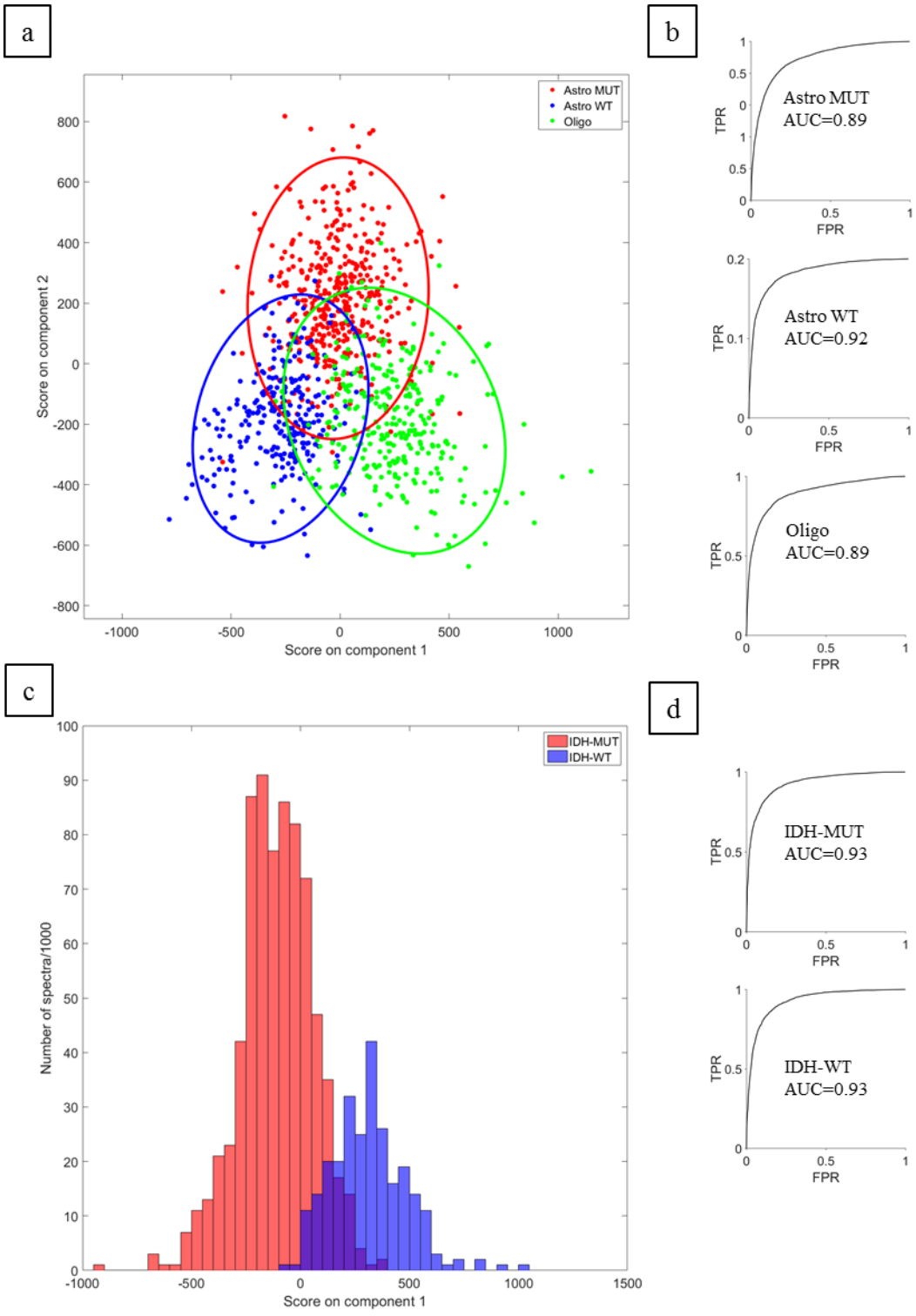


Figure 4.7 - Cryosection model performance. Cryosection LDA score plots and ROC curves for 3-group model (a, b) and 2-group model (c, d). (TPR= true positive rate; FPR= false positive rate; Astro-MUT = astroglial tumour, IDH mutant; Astro-WT = astroglial tumour, IDH wildtype; Oligo = Oligodendroglioma; IDH-MUT = IDH mutant tumour; IDH-WT= IDH wildtype tumour).

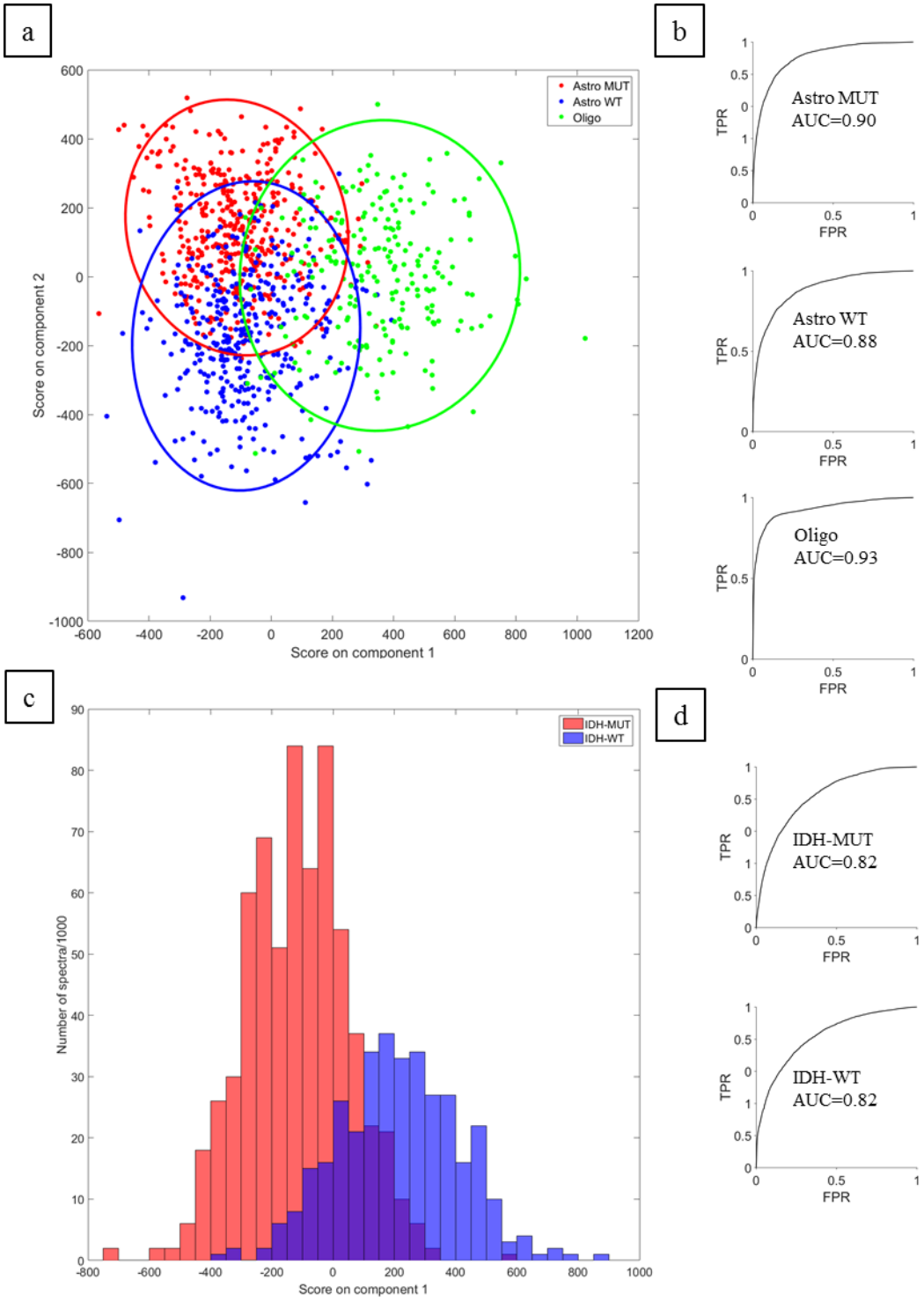


Figure 4.8 - FFPE section model performance. FFPE section LDA score plots and ROC curves for 3-group model (a, b) and 2-group model (c, d). (TPR= true positive rate; FPR= false positive rate; Astro-MUT = astroglial tumour, IDH mutant; Astro-WT = astroglial tumour, IDH wildtype; Oligo = Oligodendroglioma; IDH-MUT = IDH mutant tumour; IDH-WT= IDH wildtype tumour).

4.3.2 Mean spectra and peak assignment

For each genetic subgroup in the three-group model mean spectra and -2ndD are shown in Figure 4.9 for fresh tissue samples and Figure 4.10 for cryosections and FFPE sections. For the two-group model, mean spectra and -2ndD are shown in Figure 4.11 for the fresh tissue samples and Figure 4.12 for cryosections and FFPE sections. On all -2ndD plots the wavenumbers of prominent Raman peaks are labelled and the statistical significance of the difference between the peak intensities of each genetic subtype in that model is indicated ($p<0.001$, Wilcoxon rank sum test). Appendix C tabulates all Raman peaks according to three- and two-group models and for all sample preparation with corresponding p values. Table 4.3 summarises all the Raman peaks found in this study with corresponding tentative biomolecular assignment.

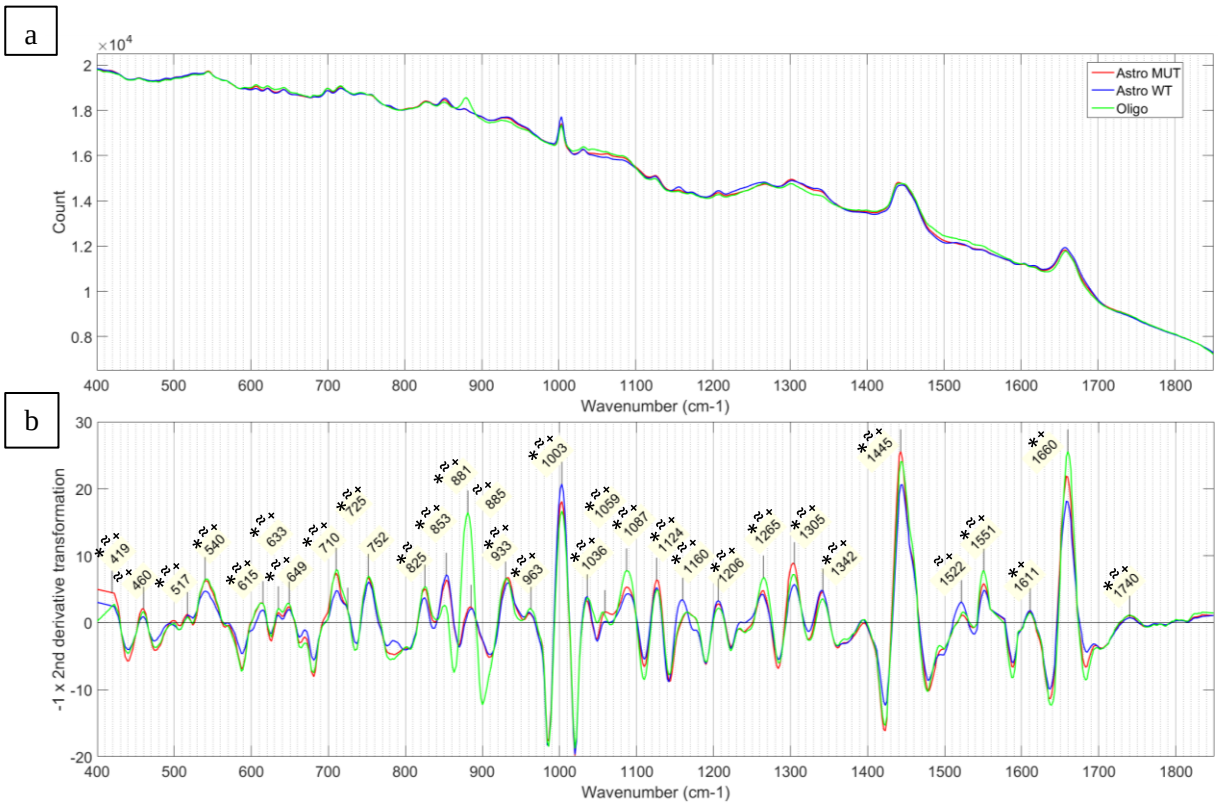


*Figure 4.9 - Fresh tissue 3-group model. a) Mean spectra and b) negative of second derivative transformation of mean spectra for each genetic subtype. The wavenumber of prominent Raman peaks are labelled in b) and the statistical significance of the difference between the peak intensities of each genetic subtype in that model is indicated ($p<0.01$) for *IDH-mutant versus IDH-wildtype; ≈ astrocytoma, IDH-mutant versus oligodendroglioma; and + astrocytoma, IDH-wildtype versus oligodendroglioma. Astro MUT = astroglial tumour, IDH mutant; Astro WT = astroglial tumour, IDH wildtype; Oligo = Oligodendroglioma.*

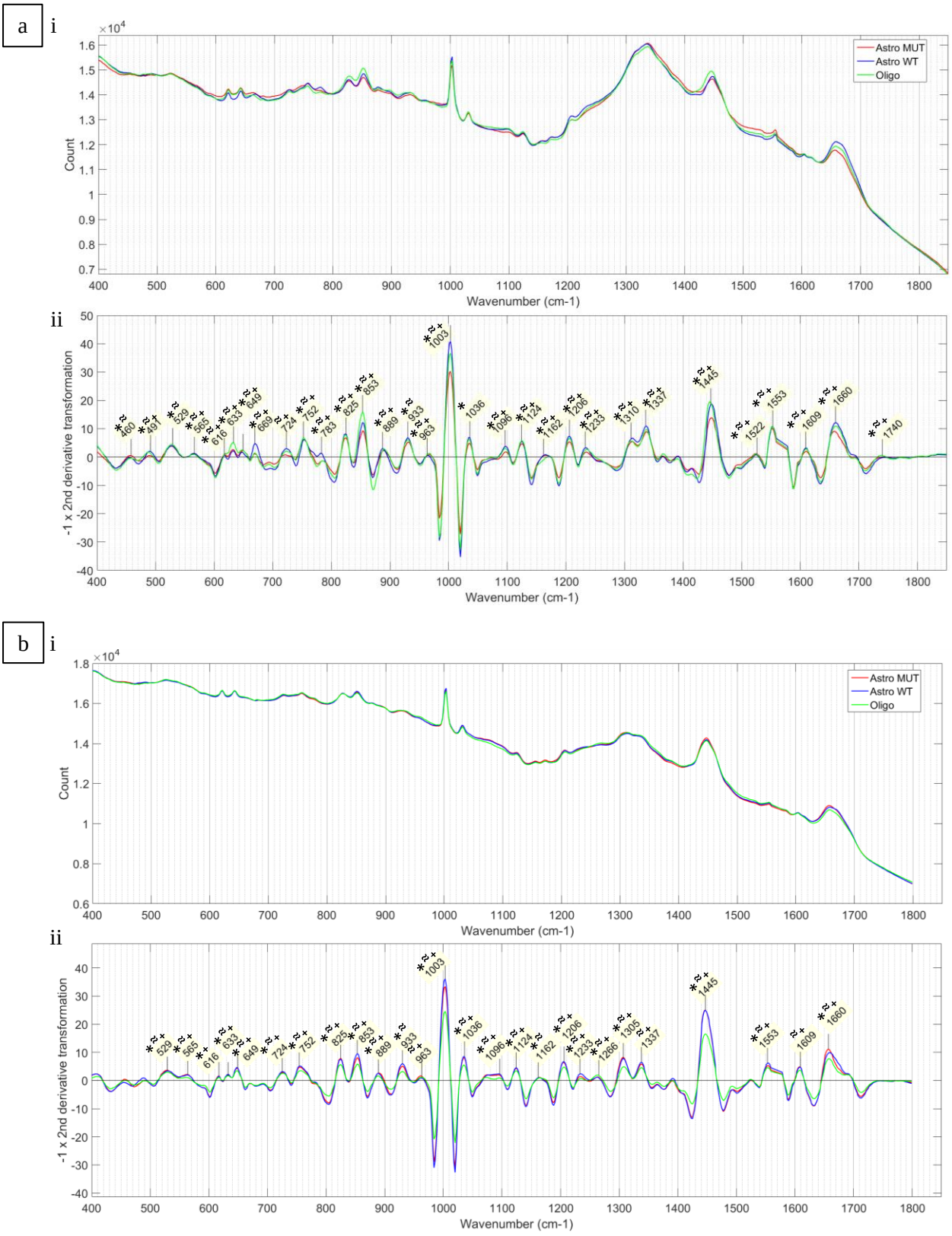


*Figure 4.10 - a) Cryosections and b) FFPE sections 3-group model. i) Mean spectra and ii) negative of second derivative transformation of mean spectra for each genetic subtype. The wavenumber of prominent Raman peaks are labelled in b) and the statistical significance of the difference between the peak intensities of each genetic subtype in that model is indicated (p <0.01) for *IDH-mutant versus IDH-wildtype; ˜astrocytoma, IDH-mutant versus oligodendroglioma; and ⁺astrocytoma, IDH-wildtype versus oligodendroglioma. (Astro MUT = astroglial tumour, IDH mutant; Astro WT = astroglial tumour, IDH wildtype; Oligo = Oligodendroglioma).*

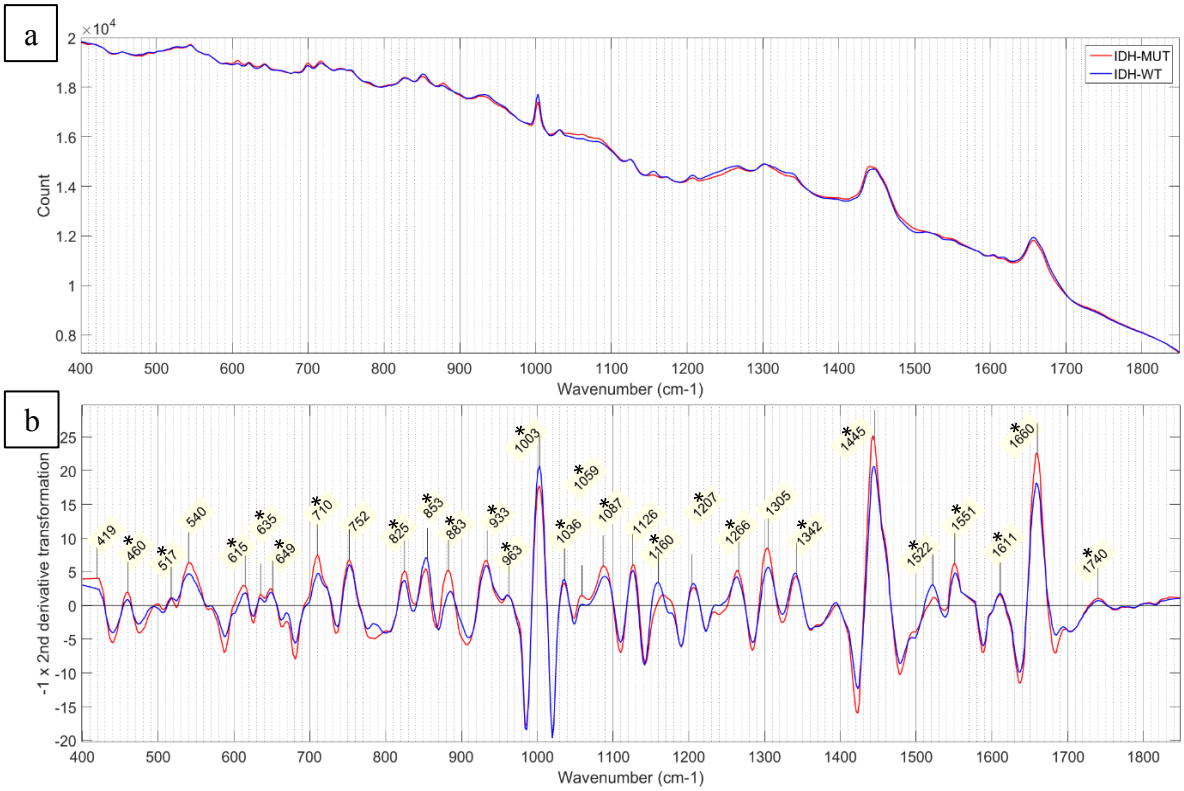


*Figure 4.11 - Fresh tissue 2-group model. a) Mean spectra and b) negative of second derivative transformation of mean spectra for each genetic subtype. The wavenumber of prominent Raman peaks are labelled in b) and the statistical significance of the difference between the peak intensities of each genetic subtype in that model is indicated (p <0.01) for *IDH-mutant versus IDH-wildtype. IDH-MUT = IDH mutant; IDH-WT = IDH wildtype*

The most dominant Raman peaks in the fresh tissue sample are seen at 1445 cm^{-1} and 1660 cm^{-1}. These are both a summation of multiple overlapping peaks and represent the ratio of protein and lipid content within the sample. There is also a characteristically sharp peak at 1003 cm^{-1} which, along with peaks at 615 cm^{-1} and 1036 cm^{-1}, represents phenylalanine. Other amino acid peaks include 460 cm^{-1} (tryptophan/tyrosine), 635 cm^{-1} and 649 cm^{-1} (tyrosine) and 540 cm^{-1} (cystein). Various peaks were assigned to classes of lipids including triglycerides/fatty acids (853 cm^{-1}, 1059 cm^{-1}, 1087 cm^{-1}, 1126 cm^{-1}, 1160 cm^{-1}, 1266 cm^{-1}, 1305 cm^{-1}, 1740 cm^{-1}) and cholesterol (419 cm^{-1}, 825 cm^{-1}). Peaks assigned to proteins include 1522 cm^{-1} and 1551 cm^{-1} (amide III) and 1660 cm^{-1} (amide I). DNA peaks were assigned at 517 cm^{-1}, 883 cm^{-1}, 963 cm^{-1}, 1207 cm^{-1} and 1342 cm^{-1}. Peaks at 710 cm^{-1} and 933 cm^{-1} most likely represent glycogen.

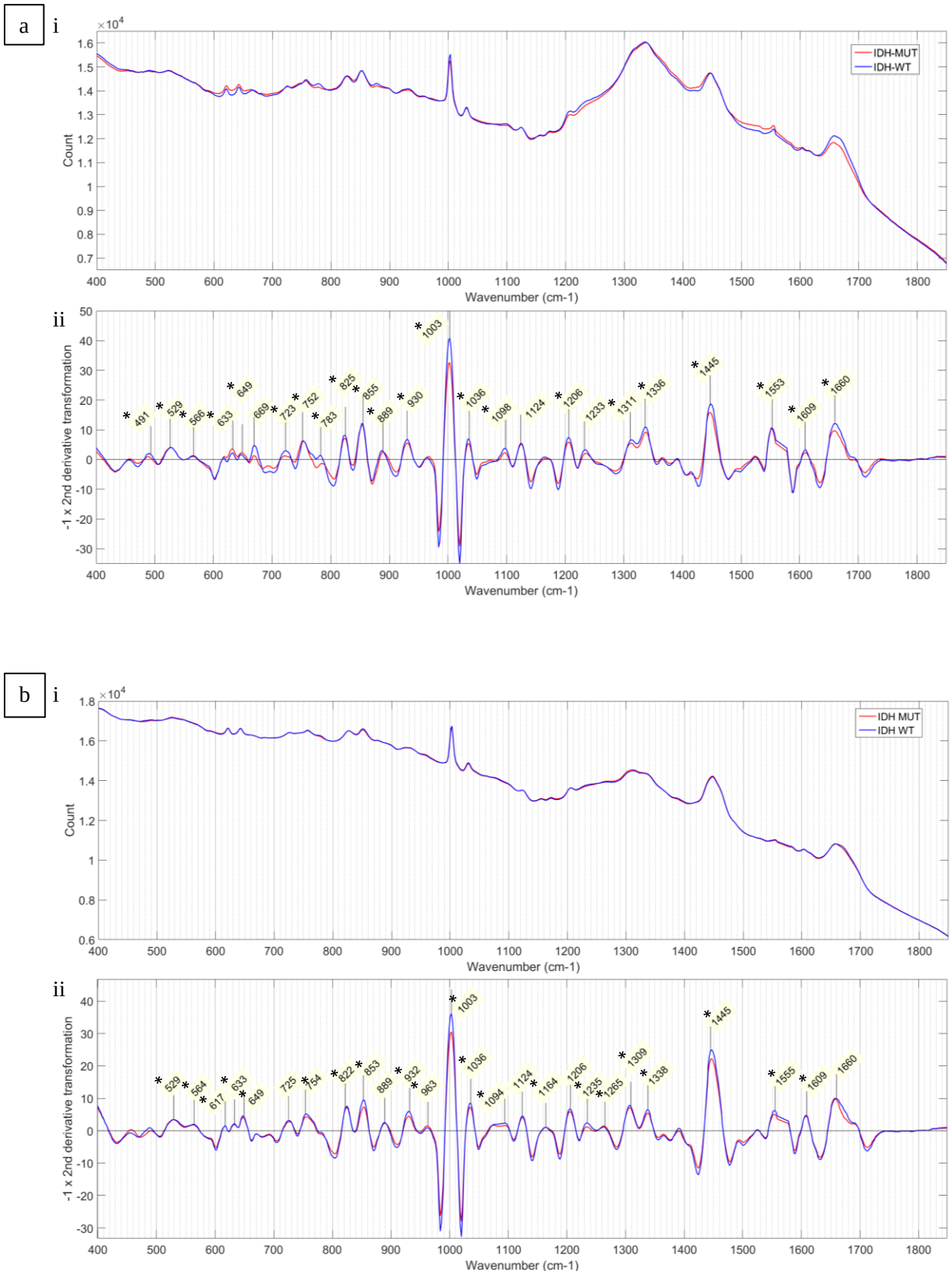


*Figure 4.12 - a) Cryosection and b) FFPE 2-group model. i) Mean spectra and ii) negative of second derivative transformation of mean spectra for each genetic subtype. The wavenumber of prominent Raman peaks are labelled in B and the statistical significance of the difference between the peak intensities of each genetic subtype in that model is indicated (p <0.01) for *IDH-mutant versus IDH-wildtype. (IDH-MUT = IDH mutant; IDH-WT = IDH wildtype).*

Wavenumber (cm^{-1})	Tentative biomolecular peak assignment	Main Biomolecule group	Secondary biomolecule group
419	Cholesterol	lipid	
460	Tryptophan/tyrosine	a. acid	
491	Glycogen (C-C skeleton), DNA, fatty acid	carb.	DNA/f. acid
517	DNA, phospholipid	DNA	lipid
529	Phospholipid, Phenylalanine	lipid	a. acid
540	Cystein, DNA, Collagen	a. acid	DNA
565	DNA, Collagen	DNA	protein
616	Phenylalanine	a. acid	
633	Tyrosine	a. acid	
649	Tyrosine	a. acid	
669	RNA + DNA, Collagen	DNA	protein
710	Glycogen	carb	
724	DNA Adenine nucleobase, RNA, Collagen	DNA	protein
752	Phospholipid, Phenylalanine, Glycogen, fatty acid	lipid	a. acid
783	DNA adenine and thymine base, RNA	DNA	
825	Cholesterol (steroid backbone)	lipid	
853	Triglyceride and fatty acid (C-COO), C-O-C of glycogen, Collagen	lipid	protein
881	DNA, phospholipid	DNA	lipid
884	DNA	DNA	
889	DNA	DNA	
930	Phospholipid (PO43-)	lipid	
933	Ring modes of glycogen, collagen, fatty acid	carb.	protein
963	DNA backbone, cholester	DNA	
1003	Phenylalanine (Collagen), DNA	a. acid	DNA
1036	Phenylalanine (Collagen)	a. acid	
1059	Triglyceride and fatty acid (C-O-C), DNA	lipid	DNA
1087	Triglyceride and fatty acid (C-O-C), Phospholipid (PO43-, C-O-C), COH deformation of glycogen	lipid	carb
1096	DNA and RNA (O-P-O),COH deformation of glycogen, collagen	DNA	carb/protein
1124	Triglyceride (C-O-C), COH deformation of glycogen, Collagen	lipid	carb/protein
1162	Fatty acid	f. Acid	
1206	DNA, Collagen	DNA	protein
1234	RNA + DNA	DNA	
1265	Triglyceride and fatty acid, phospholipid (PO43-)	lipid/f. acid	
1305	Triglyceride, phospholipid, DNA	lipid	DNA
1310	Phenylalanine	a. acid	
1337	Lipid/protein. DNA A and G nucleobase, COH and CH bend of glycogen, Collagen, fatty acid	lipid/protein	DNA
1342	DNA A and G nucleobase, RNA, collagen	DNA	protein
1445	Triglyceride-fatty acid/phospholipid (CH3, CH2 bending), collagen, cholesterol	lipid/protein	
1522	Amide III (N-H bend/C-N stretch) Priotein/DNA/RNA	Protein	
1553	Amide III (N-H bend/C-N stretch) Priotein/DNA/RNA	Protein	
1609	Phenylalanine	a. acid	

1611	Amide 1 of Protein (Collagen)	Protein	
1660	Triglyceride-fatty acid/Phospholipid (C=O ester), Amide 1 of Protein (Collagen), C-N strech of DNA, Cholesterol	lipid/protein	
1740	Triglyceride (C=O ester)	lipid	

Table 4.3 – Biomolecule peak assignment. Summary of Raman peaks found in this study on all genetic subtypes and in all tissue types with the corresponding tentative biomolecular assignment. The main biomolecular and secondary biomolecular groups represented by each peak is also summarised. (a.acid = amino acid; carb. = carbohydrate; f.acid = fatty acid)

The majority of identified peaks exhibited significant differences in intensity between the various genetic subtypes. For example, only 5 out of 30 peaks in the fresh tissue two-group model failed to exhibit a statistically significant intensity difference. When comparing IDH-mutant with IDH-wildtype intensities, in both the two- and three-group fresh tissue models, the most significant differences (lowest *p* values) were found in peaks representing lipid (853 cm^{-1}, 1059 cm^{-1}, 1087 cm^{-1}, 1266 cm^{-1}, 1445 cm^{-1} and 1660 cm^{-1}), phenylalanine (1003 cm^{-1}) and DNA (1207 cm^{-1} and 1342 cm^{-1}). The differences observed between astrocytomas, IDH-mutant and oligodendroglioma were very similar to those observed between astrocytomas, IDH-wildtype and oligodendroglioma with oligodendroglioma intensities having the most significant differences at peaks assigned to phenylalanine (1003 cm^{-1} and 1036 cm^{-1}), DNA (881 cm^{-1}, 885 cm^{-1} and 1342 cm^{-1}) and protein (1522 cm^{-1} and 1551 cm^{-1}). Peaks around 800 to 900 cm^{-1}, likely representing DNA, are the most prominent of these. In the fresh tissue model, there was a very clear dominant peak at 881 cm^{-1} (the data were reviewed to confirm that this was not an artefact). This dominant peak was not reproduced in the cryosection or FFPE section models. In general, the peak differences between oligodendroglioma and IDH-wildtype were more pronounced than between oligodendroglioma and IDH-mutant astrocytomas.

There is consistency between the dominant Raman peaks across the different tissue preparations (Appendix C). Depending on the model, between 12 and 17 peaks out of the 43 peaks identified are present in all three tissue preparations. 11 of these peaks are also present in both LN18 preparations (cryosections and FFPE sections). In general, cryosection and FFPE section spectra have more prominent DNA and amino acid peaks and less contribution from lipid and protein peaks. The cryosection spectra have a significant background in the 1200 to 1500 cm^{-1} region, which is thought to be due to water condensation on the objective lens as the cryosection dries out. There is also a strong contribution from a prominent peak at 1336 cm^{-1} (assigned to protein:lipid ratios) in this wavenumber region which is not

seen in the fresh tissue spectra. In terms of differentiating between genetic subtypes, peaks that consistently give the most significant differences between subtypes across the three tissue preparations are those assigned to tyrosine (635 and 649 cm^{-1}), phenylalanine (1003 cm^{-1} and 1036 cm^{-1}), lipid (825 cm^{-1} and 853 cm^{-1}), protein (1551 cm^{-1}), and lipid:protein ratios (1445 cm^{-1}). In the three-group models including oligodendroglioma, peaks for DNA (725 cm^{-1}, 963 cm^{-1} and 1206 cm^{-1}) are also consistent across all sample preparations.

4.3.3 Rare IDH mutations

A total of 13 rare IDH mutations were present overall, in three fresh tissue samples, six cryosections, and four FFPE sections (Table 4.4). All cases were correctly classified into the IDH mutated group (using the two-group model), with the highest probability of correct classification occurring in the fresh tissue samples, with a mean probably of 91% for predicting IDH mutant.

Tissue preparation	Histology/ grade	Rare IDH mutation confirmed with targeted genetic sequencing	Probability of predicting IDH mutation with Raman model
Fresh tissue	Astrocytoma IV	IDH1 394T	0.95
	Astrocytoma IV	IDH2 G515T	0.92
	Astrocytoma II	IDH2 G515A	0.87
Cryosections	Astrocytoma II	IDH1 R132C	0.92
	Astrocytoma II	IDH1 R132C	0.96
	Astrocytoma IV	IDH1 R394T	0.61
	Oligodengroglioma II	IDH2 R172K	0.93
	Oligodengroglioma II	IDH2 R172K	0.78
	Astrocytoma III	IDH1 R132C	0.77
FFPE sections	Astrocytoma IV	IDH1 R132C	0.57
	Astrocytoma IV	IDH1 R132G	0.82
	Oligodengroglioma II	IDH2 R172K	0.82
	Oligodengroglioma III	IDH2 R172W	0.89

Table 4.4 – Classification of rare IDH mutations. List of rare IDH mutations in found in study and probability of predicting IDH mutation with Raman spectroscopy.

4.3.4 LN18 cell line classification

All LN18 IDH-mutant cell lines stained strongly for R132H IDH mutation on IHC with zero staining of IDH-wildtype cell lines (Figure 4.13). Mean spectra and -2ndD for both the cryosections and FFPE sections are shown in Figure 4.14. The mean spectra for the LN18 cells and key Raman peaks which differentiate genetic subtypes correlate with the tissue preparation spectra. Table 4.5 gives the sensitivity,

specificity and accuracy (AUROC) using PCA-LDA for each mutant and wildtype cell lines processed as cryosections or FFPE sections.

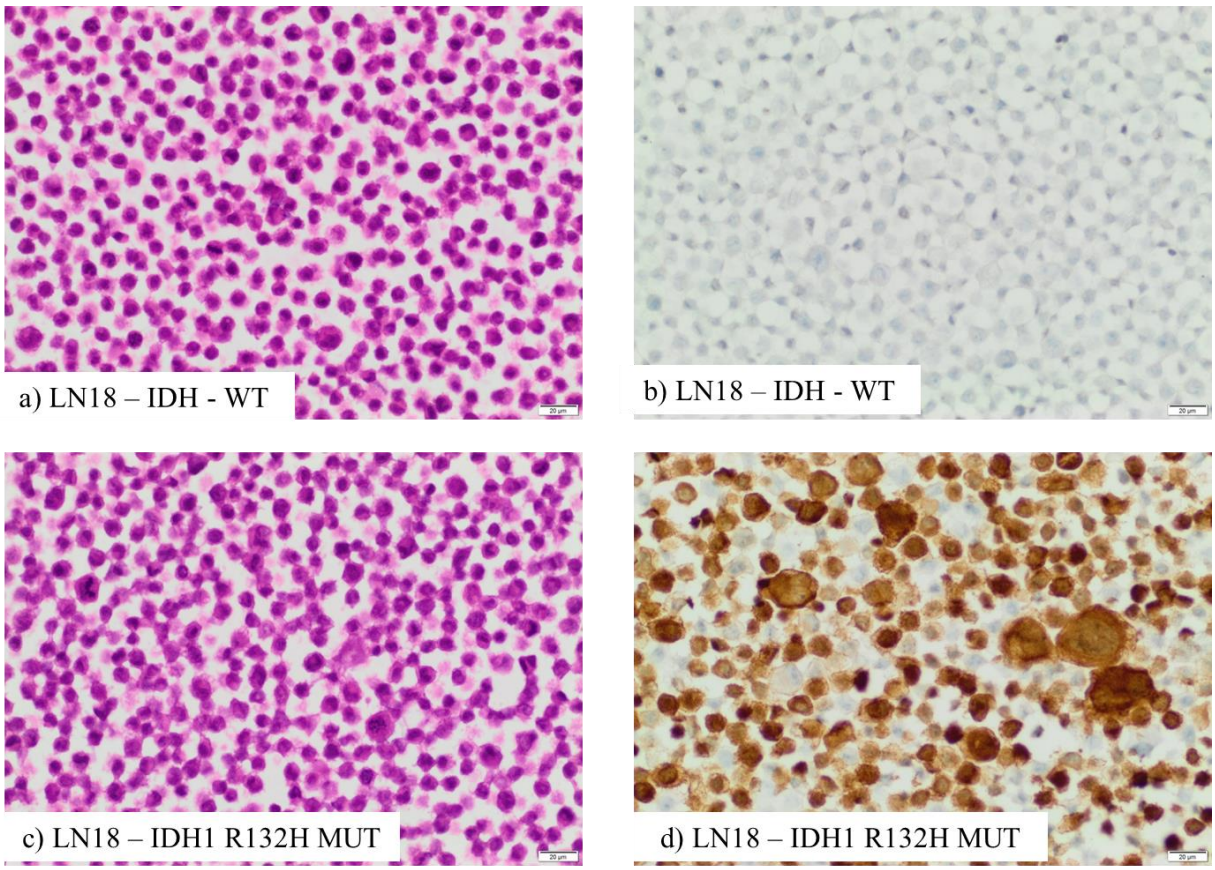


Figure 4.13 - LN18 cell line staining. H&E staining (left) and R132 IDH mutation IHC staining (right) for IDH-wildtype LN18 cell line (a and b) and IDH-mutant LN18 cell line (c and d).

Within each cell line there was excellent classification in both the cryosection and FFPE groups, with accuracy of between 0.82 and 0.99. In the FFPE group this high performance was maintained when spectra from both cell lines were combined into one model. In the cryosection group, however, the performance dropped when the cell line spectra were combined.

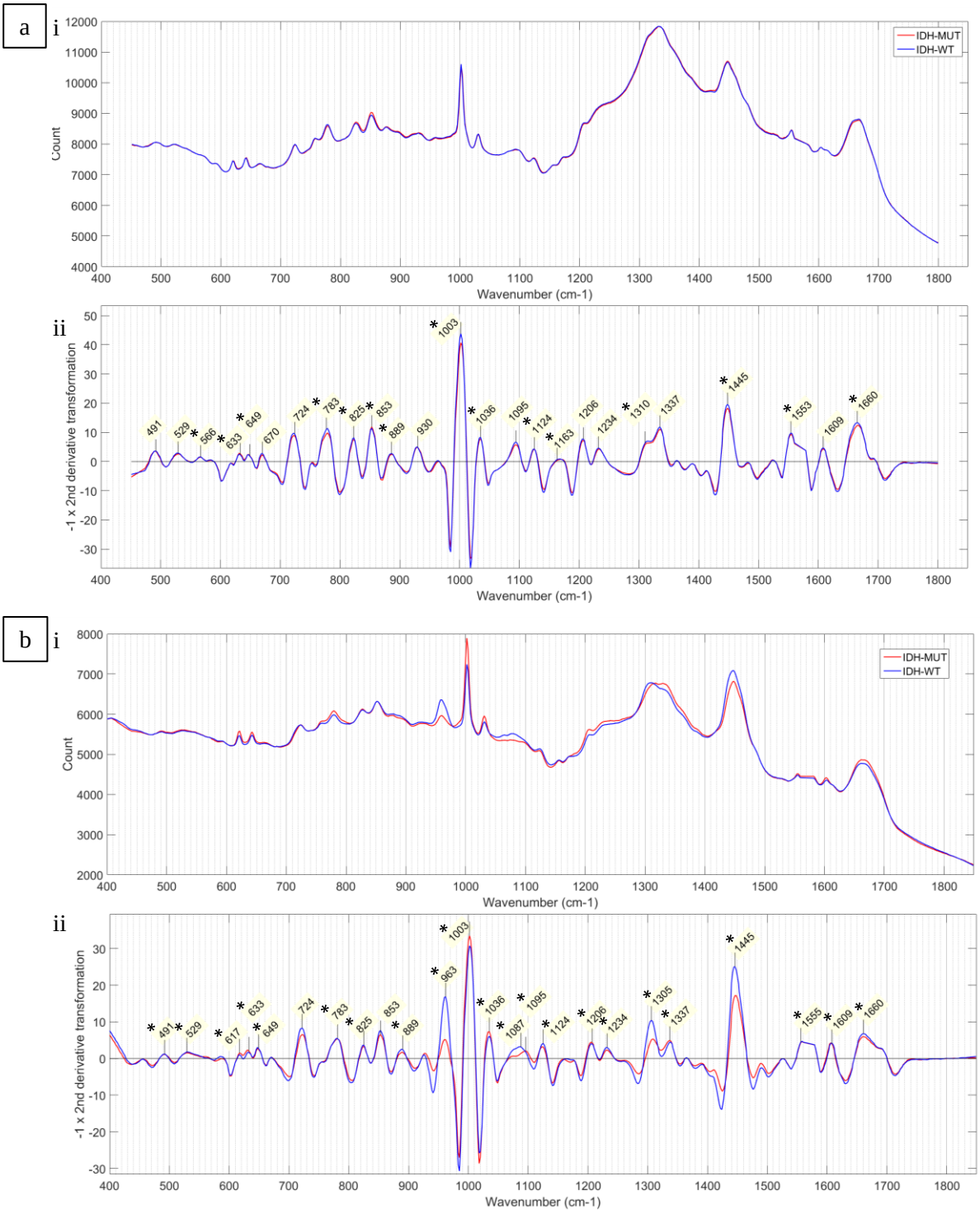


*Figure 4.14 - LN18 cell line a) Cryosection and b) FFPE 2-group model. i) Mean spectra and ii) negative of second derivative transformation of mean spectra for each genetic subtype. The wavenumber of prominent Raman peaks are labelled in a and the statistical significance of the difference between the peak intensities of each genetic subtype in that model is indicated (p <0.01) for *IDH-mutant versus IDH-wildtype. (IDH-MUT = IDH mutant; IDH-WT = IDH wildtype*

Preparation method, cell line	Sensitivity	Specificity	Accuracy
Cryosection, 1	0.85	0.70	0.82
Cryosection, 2	0.91	0.99	0.98
Cryosection, combined	0.78	0.55	0.66
FFPE, 1	0.99	0.99	0.99
FFPE, 2	0.83	0.85	0.92
FFPE combined	0.95	0.89	0.97

Table 4.5 - LN18 cell line classification performance. Sensitivity and specificity for predicting IDH mutation using PC-LDA for different sample preparation methods using spectra from individual cell line batches and combined batches.

4.4 Discussion

This study has demonstrated that Raman spectroscopy can be used to determine the most common glioma genetic subtypes using fresh tissue straight from the operating theatre, cryosections, and FFPE sections. Most significantly, the fresh tissue workflow requires no sample processing, and results are available in under 15 minutes, making this technique ideal for rapid intraoperative diagnosis.

The important role that surgery plays in the treatment of gliomas and the relationship between extent of resection and patient survival has been comprehensively reviewed in Chapter 1. Also described in Chapter 1 is how glioma surgery differs from many other types of cancer surgery, as the diffuse nature of these tumours means that the surgeon can never achieve complete tumour resection. The aim is to removal all the tumour identified on preoperative MRI scans, but it is known that tumour cells will have infiltrated beyond this boundary. The brain surrounding the tumour often controls important neurological function, so extending resection can result in neurological disability for the patient. The crucial balance for the surgeon is to maximise resection of the tumour whilst minimising the risk of disability (Figure 1.2). Surgeons vary widely as to how far they push the boundary of resection and the techniques they use to define this tumour:functional boundary. This is reflected in a wide range of 15% to 70% complete resection rates reported in the literature[71,100]. In this era of personalised medicine, the ability to perform accurate genetic subtyping of gliomas at the time of surgery will allow the surgeon to adapt their surgical strategy with some knowledge of the likely prognostic and treatment potential for the lesions. This study has shown that Raman spectroscopy can accurately identify the three most common and clinically relevant glioma groups: astrocytomas, IDH-mutant; astrocytomas, IDH-wildtype; and oligodendrogliomas, IDH-mutant and 1p/19q-codeleted. The importance of IDH mutation and 1p/19q-

codeletion has been summarised in Chapter 1. More relevant to intraoperative decision making, Wijnenga *et al.*, amongst others[97,98], have examined the impact on overall survival of the extent of surgical resection, measured as volume of residual tumour seen on postoperative contrast MRI, in molecularly stratified low grade gliomas (WHO II astrocytomas and oligodendrogliomas)[100]. They reported that the effect of less tumour residual volume on overall survival was much greater in IDH-mutant astrocytomas compared to oligodendrogliomas. In the IDH-mutant astrocytoma group, even a small amount of residual tumour of 0.1 cm^3 impacted negatively on survival compared to 0 cm^3 residual. This was not seen in the oligodendroglioma group, with a small residual having no significant effect on overall survival. Consequently, pushing the boundaries of surgical resection when there is a possible increased risk of disability might be in the patient's best interest if they have an IDH-mutant astrocytoma, but would not be in their interest if they have an oligodendroglioma: intraoperative information of the glioma genetic subtyping is therefore essential. As more is learnt about glioma subtypes and novel therapies are tailored for their treatment, including insertion of therapy agents at the time of surgery, the need for accurate genetic subtype information at the time of surgery will increase. A collaboration with our group[251], and other groups, have shown that preoperative non-invasive magnetic resonnance spectroscopy (MRS) at 3- or 7-Tesla may provide the neurosurgeon with information on IDH status of diffuse gliomas[252] however, to our knowledge, this currently does not allow distinction between IDH-mutant astrocytomas and oligodendrogliomas.

The major advantage of Raman spectroscopy compared with conventional techniques is the lack of sample preparation required. The protocol developed here could easily be replicated by anyone in the operating theatre once fully automated Raman analysis was integrated into the workflow. The fresh tissue samples simply require the biopsy to be smeared into a grooved slide and then placed on the Raman microscope stage (Figure 4.2). In the experimental protocol the sample was split into two sections by the neuropathologist prior to Raman analysis, as adjacent tissue H&E staining was essential in order to establish the quality of the biopsy and tumour content, and spectral collection was supervised by a researcher. These steps would not be required as part of a Raman diagnostic pathway, particularly with the straightforward automation of the process which, at the press of a button, would perform spectrometer quality control, laser autofocusing, spectral acquisition, spectral quality check, pre-processing and

classification. This sets Raman spectroscopy apart from other techniques that have been used to acquire intraoperative genetic subtyping, such as polymerase chain reaction-based rapid genotyping assays[253,254]. More traditional techniques can give genetic stratification within the timeframe of an operation, and are more accurate than Raman spectroscopy, but they all require complex laboratory-based sample processing and analysis by qualified laboratory technicians. They have significant consumables costs per test, compared to the minimal cost of the grooved stainless steel slide used for Raman analysis, which after the procedure could be sterilised, reused, and become part of the standard neurosurgical instrument set. Others have reported the use of Raman and ATR-FTIR spectroscopy to distinguish IDH-mutant versus IDH-wildtype tumours, but have used snap frozen sections prior to spectroscopy, which still requires laboratory-based sample preparation[194,212]. These vibrational spectroscopy studies have not attempted to distinguish oligodendrogliomas as a separate group, which is essential for clinical decision making. As outlined in Chapter 1, Raman spectroscopy has also been used intraoperatively in the form of a probe device to attempt to differentiate tumour boundaries[190]. A probe device is essential for boundary detection, and is explored further in Chapter 6, but adds no benefit to the aim of this study: as the tumour is being removed by the surgeon anyway, in- or near-operating theatre analysis is sufficient and allows for more sensitive Raman spectrometry than would be achievable via a fibre optic probe setup.

A barrier to the translation of Raman spectroscopy into the clinical setting is that clinicians are sceptical about the reliance on a change in spectral patterns, rather than a specific biomarker measurement, to define different pathological entities. The spectral characteristics of glioma found in this study add to a body of published evidence which consistently defines similar key spectral features of glioma tissue[174,176,182,183,194,196]. Moreover, the most significant spectral differences seen between genetic subgroups are broadly consistent with the understanding of the biology of these tumours. IDH-mutant and IDH-wildtype gliomas are known to have significant differences in lipid metabolism[247,255-257], collagen maturation[258] and DNA methylation[13,259] which correlates with the most differentiating Raman peaks assigned to lipids, phenylalanine (involved in collagen synthesis) and DNA. The 881 cm^{-1} peak is markedly different in the fresh tissue oligodendroglioma spectrum. The explanation for this is not entirely clear; however, whilst superficially similar (regarding high levels of 2HG as a result of an IDH mutation),

IDH-mutant astrocytomas and oligodendrogliomas are otherwise clinically and molecularly distinct. It is also known that certain metabolic features are associate with the 1p/19q codeletion rather than IDH mutation (e.g. cystathionine[260]). As altered phospholipid metabolism is known to occur in subsets of glioma[261], it cannot be ruled out that the increased peak at 881 cm^{-1} reflects these metabolic changes rather than an increased amount of DNA. The presence of other peaks in the oligodendroglioma mean spectrum, which possibly represent phospholipid (518 cm^{-1}, 1089 cm^{-1} and 1300 cm^{-1}), support this theory. Indeed, there have been previous reports of the importance of phospholipid Raman peaks in differentiating glioma tissue from normal brain tissue[186]. The lack of phospholipid peak dominance in the cryosection and FFPE section models does not support this hypothesis, however. It is important to note that the present study did not aim to investigate proteomic or metabolomic profiles of different genetic subtypes – there are more suitable techniques to do this than Raman spectroscopy. However, this correlation with known tumour biology does reinforce and validate the biological origin of the spectra.

There is considerable consistency of Raman spectral peaks across all of the different sample preparation methods and cell lines employed in the present study. As might be expected, the performance of the classification model decreases as the tissue becomes more manipulated by the sample preparation, with fresh tissue models performing better than cryosections, which in turn classifies better than FFPE sections. Spectra of cryosections and FFPE sections have fewer contributions from proteins and lipid, differentiating genetic subtypes more on DNA and amino acid peaks. This is consistent with the known effect of protein denaturing due to freezing, cross linking of proteins as part of formalin fixation, and the loss of lipid due to paraffin embedding. In the protocol development it was found that cryosections could tolerate much higher laser exposure without CCD saturation or apparent deterioration in the spectral quality than both fresh tissue and FFPE sections. Indeed, the cryosection exposure time was up to 10 times greater than that used to measure fresh tissue samples. It is perhaps surprising, therefore, that the cryosection model performance is not greater given the better SNR achieved. This suggests that there is a definite effect of sample freezing on spectral quality and possibly highlights the importance of large sample thickness giving fresh tissue an advantage.

The LN18 cell line classification was very accurate within individual cell line preparations. In the cryosection preparations, the performance dropped off when the model includes spectra from both cell lines. This may be the result of large intragroup variability, as there are only two groups of each mutation and wildtype cells in the model. The performance would be likely to increase when the number of cell line batches is much greater, causing a decrease in intragroup variability and an increase in the intergroup variability. The strong performance in classifying these cell lines is consistent with a number of reports of the successful application of Raman as a cell sorting technique, which is an area of active research[262-264].

The present study has also shown that Raman spectroscopy has application in the neuropathology and research laboratory settings, offering rapid genetic subtyping of gliomas using cryosections, FFPE sections and glioma cell lines. The minimal sample preparation and rapid analysis advantages of Raman spectroscopy are less critical within a laboratory setting. However, routine IHC testing for IDH mutation is only sensitive to the common IDH1 mutation, whereas Raman spectroscopy measures the downstream biological consequences of the mutation, most likely associated with the intratumour accumulation of 2-hydroxyglutarate, which is common to all glioma-associaed IDH1 and IDH2 mutations[265]. In our study the IHC test for the common IDH1R132H mutation was, as expected, negative in those cases harbouring rare IDH1 or IDH2 mutations (Table 3). This amounted to a significant proportion of our cohort (5%) and all cases were correctly classified as IDH-mutant in the Raman analysis. High-throughput assays based on next-generation sequencing or microarray methylation chip techniques are likely to replace single marker assessment by IHC, Sanger sequencing or FISH, but these are not currently widely available and have a five to seven day turnaround time. Raman analysis may provide a faster and lower-cost alternative, certainly for initial tissue screening in environments where access to molecular diagnostics may not be available.

In summary, this study has demonstrated the potential for Raman spectroscopy to be employed in a straightforward way within the neurosurgical and neuropathological glioma diagnostic pathway to provide genetic subtyping of gliomas within 15 minutes of the biopsy. The resulting data gives the surgeon the opportunity to tailor the surgical strategy to the patient's specific tumour genetics. Despite this being the largest study of its kind, the small sample size is a limitation. Further multicentre and larger

studies need to be undertaken to expand the training data set and to evaluate the effect of inter-user and inter-machine variability on the predictive model. Finally, the predictive model needs to be tested with a prospective independent cohort of patients to robustly validate the technique.

Chapter 5. Development of a laboratory-based fluorescence setup, fluorescence standard and comparison of operative microscope and endoscope fluorescence detection thresholds

5.1 Introduction

The use of 5ALA-induced fluorescence-guided surgery is considered to be the gold standard for the resection of HGG. As detailed in Chapter 1, there is excellent correlation between visual fluorescence and the presence of tumour (Section 1.5.2). However, it is also known that tumour exists beyond the visual fluorescence boundary, giving visual fluorescence a low negative predictive value. Chapter 1 reports on the research attempting to increase detection of the 5ALA-induced (or porphyrin IX (PpIX)) fluorescence, including the development of intraoperative fluorescence probes (Section 1.7). Our group are developing novel methods of enhancing the detection of PpIX fluorescence and this Chapter reports on the development of a fluorescence standard and laboratory-based fluorescence setup. The Chapter then outlines the use of this new setup to test the hypothesis that greater visual fluorescence is detected using the operative endoscope compared with the operative microscope.

Access to the operating theatre is limited due to frequency of use for surgeries and restrictions due to sterility and infection control protocols. Intraoperative testing of equipment is possible but needs to be justified as this results in lengthening of the surgical procedure and a longer period of anaesthetisation for the patient. The operative microscope and endoscope are also difficult to use for testing as they are constantly required for surgical procedures. In the case of the microscope, transportation out of theatres, even to a laboratory on the same level in the same building, is very difficult and not recommended by the manufacturer. There is therefore a need to develop a laboratory-based fluorescence measurement setup to allow for testing of any new device or instrument outside of the theatre setting prior testing in the intraoperative environment. The development of a fluorescence spectrometry probe-based system was also important to allow for future comparison of this technology with a Raman probe-based system. The aim of this study was to develop and validate a laboratory-based fluorescence measurement setup

and a fluorescence standard that could be used as a model for the 5ALA-induced fluorescence guided surgery carried out in the neurosurgery operating theatre. Once established, this setup was used to compare the microscope and endoscope fluorescence detection thresholds to test the hypothesis that the endoscope can detect fluorescence to a lower concentration of PpIX compared to the microscope.

The objectives were to:

1. characterise the fluorescence light sources used in neurosurgical operating theatres and to correlate these with a similar light source to be used in the laboratory setup;
2. measure PpIX fluorescence using the laboratory-based setup and assess its suitability to be used as a laboratory standard for fluorescence measurements;
3. develop an alternative fluorescence standard to act as a surrogate of PpIX and to correlate this standard with PpIX fluorescence;
4. use the developed experimental setup to compare fluorescence detection thresholds using operative microscope versus endoscope light sources.

5.2 General methods

Throughout these experiments, spectra were acquired using the collection fibre of a reflectance probe (product code: RP22. Thorlabs Ltd, UK) coupled to a high resolution spectrometer (product code: HR4000. Ocean Optics Inc., USA). Blackout sheeting was used to limit background light, and a background spectrum, recorded with the same exposure time as the spectral measurements, was always taken prior to spectral measurements and subtracted prior to analysis. In general, multiple spectra were collected and averaged for each experimental setup. All spectra were recorded using OceanView v1.6.3 spectroscopy software (Ocean Optics, USA) and processed using Excel 2016 (Microsoft, USA) and Matlab R2015b (MathWorks, USA). The exposure time was varied in order for the amount of light collected within the acquisition to remain within the range of the spectrometer. As part of the processing, spectra were scaled to account for any differences in exposure times, allowing the spectra to be compared directly. The units of intensity are therefore arbitrary, but comparable throughout each set of experiments.

Simple linear regression using robust (bisquare weights) least squares[xvi] fitting was used to plot regression lines, and the corresponding coefficient of determination (R^2)[xvii] was calculated. All linear regression was performed using the Matlab Curve Fitting Toolbox. Non-linear spectral curve fitting models were built using the 'Solver' Microsoft Excel add-in[266].

5.3 Characterisation of light sources

Before carrying out the spectral measurements on PpIX, a set of experiments were performed in order to characterise the blue light sources from the operative microscope and endoscope used in fluorescence guided surgery and to compare these with a laboratory light source chosen to act as a substitute for the microscope and endoscope in later measurements.

5.3.1 General light source description

The operative microscope is the Zeiss OPMI® Pentero® 900 (Carl Zeiss, USA) neurosurgical microscope equipped with "BLUE™ 400" fluorescence capabilities to allow for fluorescence-guided tumour resection. The working distance for the microscope can be changed between 200 mm to 404 mm and the light intensity, for both blue and white sources, can be varied between 0 to 100% in 1% increments. The microscope has a built-in default function which automatically changes the total light intensity as the focal distance changes to maintain a contestant intensity per unit area over the illuminated surface. This is more important for the higher-powered white light source and is a safety feature to prevent tissue drying. In practice, surgeons tend to turn off this feature in order to achieve maximum possible light intensity over the operative field, and take other measures, such as irrigation, to present tissue drying. For this reason, all experiments are carried out with this function turned off. In general, the working distance was kept to the minimum of 200 mm to mimic the maximum illumination possible during surgery.

The endoscope is a Karl Storz Hopkins® 0°, 4 mm diameter, endoscope (Karl STORZ Endoscopy (UK) Ltd., UK) with both blue and white light illumination. The endoscope camera is a Karl Storz IMAGE1

[xvi] Robust least squares fitting minimises the influence of outliers on the least squares fitting. Bisquare minimises a weighted sum of squares with the weighting given to each data point depending on how far the point is from the fitted line. Points near the line get full weight. Points farther from the line get reduced weight.

[xvii] For simple linear regression R^2 is defined as the proportion of the total variation in the dependent variable that is explained by the model.

H3-Z Three-Chip FULL HD camera (Karl STORZ Endoscopy (UK) Ltd., UK). The blue light source is supplied by a 300 Watt Xenon bulb (D-Light C Light Source, Karl STORZ Endoscopy (UK) Lrd., UK). The light intensity for both blue and white light sources can be varied between 20% and 100% in 1% increments. The working distance of the endoscope can be any distance from the tissue surface to approximately 10 cm away. The endoscope tip is constantly moved throughout the operation, following the surgeon's instruments, and a practical working distance range would be between 0.5 mm and 40 mm, with a working distance of ~30 mm for much of the operation. Endoscopes with tip angles of 0°, i.e. looking straight ahead, and 30°, looking at an angle, are used by surgeons. Preliminary testing confirmed that the scope angulation did not have a significant effect on the light intensity or the area illuminated, and therefore the 0° scope was used for all of the following experiments in order to keep the optical geometry the same as that for the microscope.

To develop a laboratory-based model, a separate light source was chosen to match the wavelength range of the surgical devices as well as possible. A Thorlabs fibre-coupled light emitting diode (LED) with a nominal wavelength of 420 nm (product code: M420F2. Thorlabs Ltd, UK) was selected. The light source was controlled by an LED drive unit (T-cube LED Drive, product code: LEDD1B. Thorlabs Ltd, UK), which provided control over the output power of the light source on an arbitrary scale of 0 to 6 in 0.5 increments.

Figure 5.1 shows the spectra for each light source, normalised to a peak intensity of 1 for ease of comparison. All light sources have similar wavelength ranges, although the endoscope has a broader spectral range. Maximum intensity was at 419 nm for the Thorlabs LED, 423 nm for the endoscope and 421 nm for the microscope. All of these peak wavelengths fall within the wavelength range of PpIX absorption (375 - 440 nm[123,267]), as would be expected.

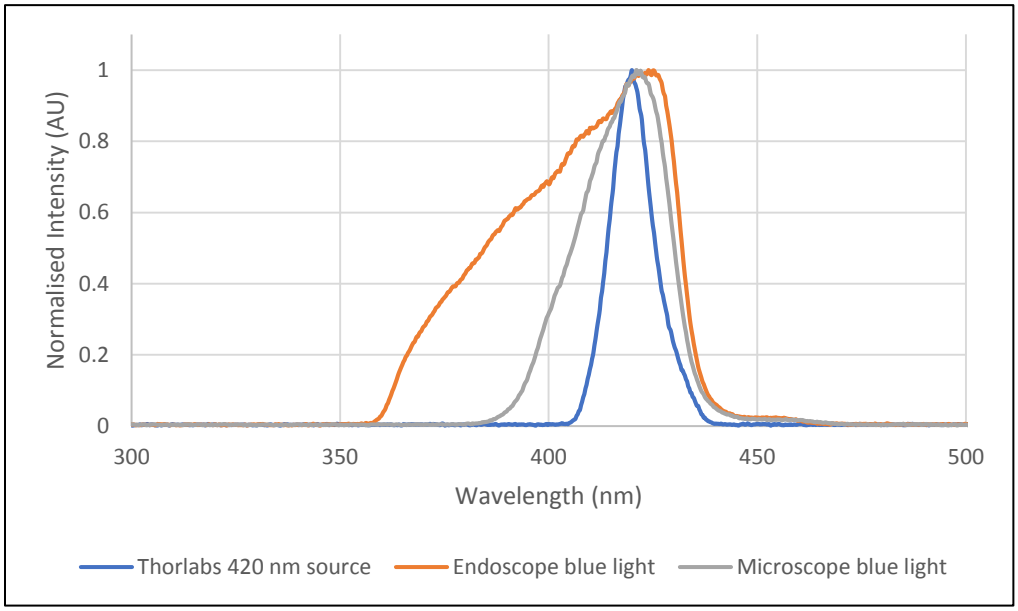


Figure 5.1- Normalised spectra for Thorlabs 420 nm source, endoscope blue light source and microscope blue light source (AU= arbitrary units)

5.3.2 Light source power calibration

The amount of PpIX fluorescence depends on the light intensity absorbed by the sample. The intensity, *I*, at the sample surface is related to the distance *r* from the light source to the surface by the inverse square law, which states:

$$I \propto \frac{1}{r^2} \qquad [3]$$

For a given light source power, *P*, assuming a spherical distribution of light over area *A*, where $A = 4\pi r^2$, the intensity at the surface is given by:

$$I = \frac{P}{A} = \frac{P}{4\pi r^2} \qquad [4]$$

From these equations it is clear that for a given light source, the distance from the surface has a significant effect on the illumination intensity at the surface. In the case of the microscope the minimum working distance is 200 mm, and therefore the maximum light intensity possible at the illuminated surface is obtained when the microscope is set at this distance. The endoscope differs in this regard as the working distance can be changed arbitrarily, and the maximum intensity is obtained when the endoscope tip is as close to the surface a possible.

The aim of this experiment was to perform a direct comparison between the illuminating power of each light source under normal operating conditions. This will allow comparisons to be made between the

amount of fluorescence seen using the microscope and the endoscope, and to correlate this with fluorescence excited with the Thorlabs LED light source, and therefore to enable laboratory-based experiments using the LED to be correlated with the endoscope and microscope situations. The calibration was carried out with the microscope and endoscope in a vertical position (90° to the horizontal) and the Thorlabs LED light beam at 45° to the horizontal. The working distance was fixed at 200 mm for the microscope. Several working distances were used for the endoscope, ranging from 5 mm to 40 mm, and for the Thorlabs LED working distances of 12 mm, 20 mm and 30 mm were used[xviii]. The light intensity settings for each source were varied, and a photodiode (silicon photodiode (product code DET210), Thorlabs Ltd, UK) was used to measure the relative photon flux (photons per unit time per unit area[xix]) for each light source setup. Multiple readings were taken for every setup and a mean calculated.

Table 5.1 shows the relative photon flux[xix], represented by photodiode signal, as the endoscope working distance is varied. The relationship agrees with the inverse square law, with the measured photon flux being inversely proportional to the square of the working distance, *r* (Figure 5.2). The photodiode

Working distance (mm)	Photodiode signal (V)
5	17.00
10	7.03
15	3.30
20	2.12
30	1.18
40	0.72

Table 5.1 - Change in photodiode signal as a function of endoscope working distance

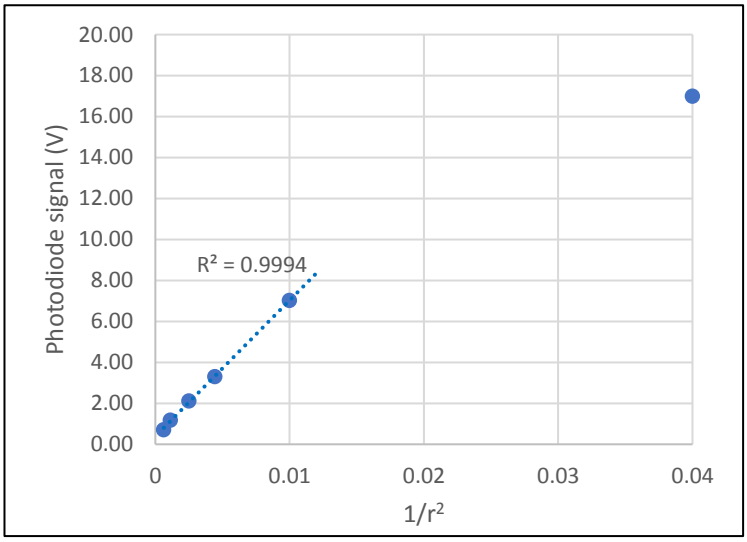


Figure 5.2 Inverse relationship between photon flux (represented by photodiode signal) and the square of the endoscope working distance (r). Dotted line represents fitted linear regression model with corresponding R^2 value (omitting highest value of $1/r^2$ of 0.04 which is thought to be an outlier due to photodiode saturation).

[xviii] The minimum working distances for the Thorlabs LED was based on the minimum distance used in later experiments using this light source.

[xix] Photon flux is defined as the number of photons per second per unit area. Here it is represented by the voltage output of a photodiode which measures photons per unit area per second and converts this to electrons.

becomes saturated at a working distance of less than 10 mm and therefore the photodiode signal at the working distance of 5 mm is excluded from the linear regression line of best fit.

Figure 5.3 illustrates the photon flux for each light source as a function of intensity settings. For all light sources there is an increase in the photon flux as the intensity setting is increased. This is a linear relationship for the microscope and endoscope light sources, but not the Thorlabs LED.

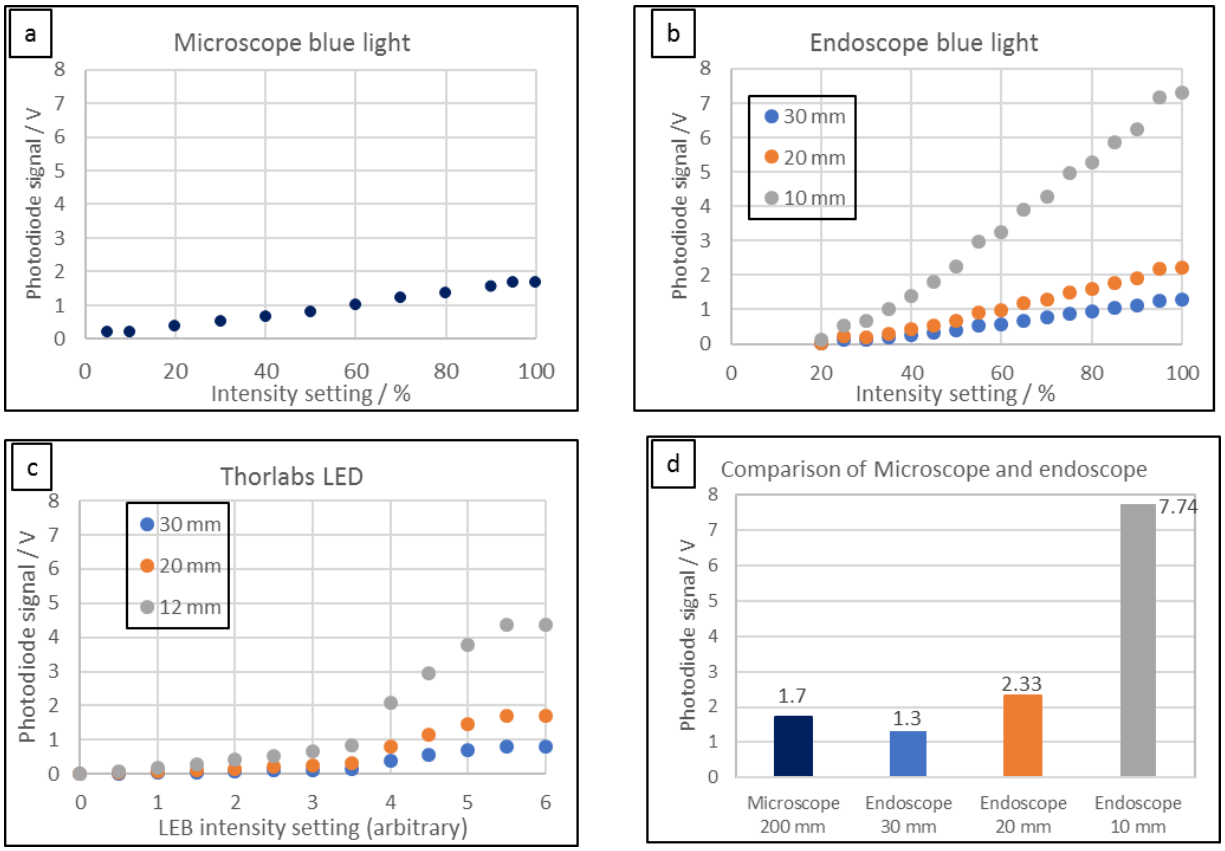


Figure 5.3 - Photodiode voltage versus light intensity setting for a) microscope at 20 mm working distance, b) endoscope at 30 mm, 20 mm and 10 mm working distances and c) Thorlabs LED at 30 mm, 20 mm and 12 mm working distances. d) Comparison of photodiode signal when light source at maximum light intensity for microscope at 200 mm working distance and endoscope at 30 mm, 20 mm and 10 mm working distances.

When comparing the light sources at their maximum intensity settings, the microscope photon flux is slightly greater than the endoscope's at 30 mm working distance (1.7 V versus 1.2 V; Figure 5.3d). However, as the endoscope working distance is reduced there is a sharp increase in the photon flux, which substantially outperforms the maximum that can be achieved by the microscope. This will be revisited when the microscope and endoscope PpIX fluorescence is compared in Section 5.8.

To use the Thorlabs LED as a surrogate for the operative light sources in further experiments we need to correlate the intensity of the Thorlabs light source with the equivalent operative light sources. This will

obviously depend on the experimental setup employed, but assuming the experimental setup described above, with the Thorlabs set to maximum power intensity and at 45° to the horizontal, correlation factors can be calculated in a straightforward way. The Thorlabs LED source intensity, I_{LED}, can be correlated with that of an operative light source, I_S, using the expression:

$$I_{LED}\, F = I_S \qquad [5]$$

where F is a constant for a given working distance and light source. Table 5.2 shows the values of F for each operative light source and working distance combination.

Operative light source and working distance	**Multiplication factor (F) for different Thorlab LED working distances**		
	***12** mm working distance*	***20** mm working distance*	***30** mm working distance*
Microscope at 200 mm	0.39	1.00	2.09
Endoscope at 30 mm	0.30	0.77	1.60
Endoscope at 20 mm	0.51	1.31	2.74
Endoscope at 10 mm	1.67	4.31	8.99

Table 5.2 - Multiplication factor (F) to correlate Thorlab LEB light source (LED) with surgical device light source (Is)

5.3.3 Summary

The intensities of the two operative light sources have been characterised for different intensity settings and working distances. The endoscope provides greater light intensity than the microscope when used at a working distance of 20 mm or less. This has practical implications for the surgeon, who must keep the endoscope within this working distance range in order to achieve an illumination benefit over the microscope. The Thorlabs LED light source was also characterised, and found to provide an acceptable substitute for the operative light sources when undertaking laboratory-based experiments.

5.4 PpIX fluorescence and photochemical degradation

Having characterised the various operative and lab-based light sources, the next step was to measure PpIX fluorescence in the laboratory setting, and to evaluate the suitability of using PpIX in solution as a standard to be used to compare the fluorescence emission following excitation excited by different light

sources. In order to use a standard solution of a substance to measure and compare fluorescence between difference light sources it is essential that there is no change in the fluorescence spectrum of that substance, in terms of both intensity and fluorescence peaks, as a result of exposure to a light source. Time series experiments were carried out to assess if there was a change in the fluorescence spectrum as a function of exposure time. In this section the fluorescence experimental setup is described and then this is used to carry out the PpIX time series.

5.4.1 Fluorescence experimental setup

Figure 5.4 shows a schematic of the experimental setup used for all fluorescence measurements employing the Thorlabs LED light source. The light source was connected to the light source input of a

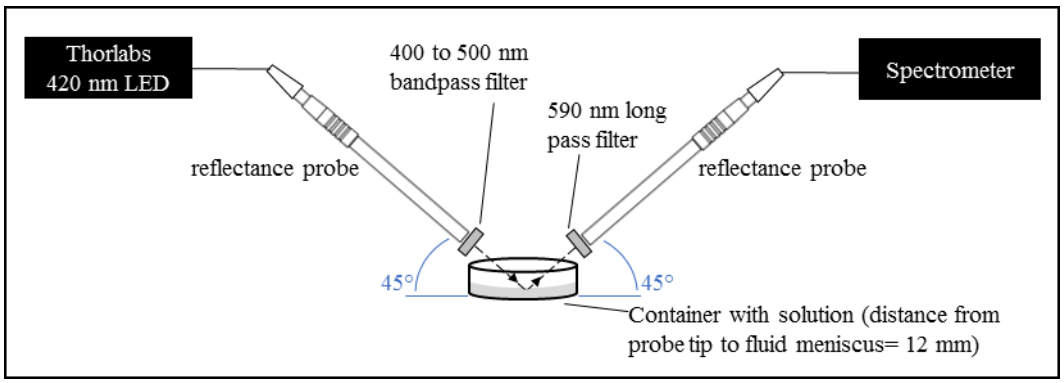


Figure 5.4 Fluorescence experimental setup with Thorlabs LED light source

reflectance probe mounted at 45° to the sample surface. It was observed during some background light measurements, and during some other preliminary work carried out using high-sensitivity cameras, that the Thorlabs LED light source has a long-wavelength 'tail', and emits a small amount of light in the fluorescence range of 590 to 700 nm. Since this would interfere with the PpIX fluorescence measurements in this wavelength region, a 400 to 500 nm bandpass filter (Custom made, Galvoptics, UK) was placed in the light source path to remove any light above the desired excitation range. Fluorescence was measured by the Ocean Optics spectrometer via the spectrometer output of a second reflectance probe mounted at 45° to the sample surface. A 590 nm long pass filter (Thorlabs, UK) was placed in front of the collection probe to filter out the excitation light. (See section 5.2 for further details of the reflectance probe and spectrometer). While in principle the measurements could have been made using a single reflectance probe, the requirement for wavelength filtering of both the illumination light

and reflectance signal made it much more straightforward to employ two probes in the arrangement shown.

In the initial time series experiments the porphyrin solution (details of the solution used will be provided in Section 5.4.2) was pipetted into a white measuring boat, which was placed directly under the light source (Figure 5.5a). While this precluded any reflectance or scattering effects introduced by a cover, it was later found that covered and sealed samples were much more stable under experimental conditions. Consequently, in later experiments a 30 mm diameter petri dish with a lid was used (Nunclon™ microplates, Thermo Scientific, USA). The lid was sealed around the sides of the container using Parafilm® (Bemis Company Inc., USA), ensuring that no Parafilm covered the top surface of the lid through which the sample would be illuminated (Figure 5.5a). The use of a lid was felt necessary to prevent evaporation of the solution and also, as the aim was to create a stable standard for use in the laboratory, it was felt that these sealed petri dishes could be stored and reused as fluorescence standards as long as the solution inside did not degrade over time. The effect of the lid on spectral quality was investigated using a solution of DCM laser dye (introduced later). As shown in Figure 5.5b, there was a slight decrease in the intensity with the addition of the lid, but there was no change in the spectral shape or wavelength of the absorption maximum (616 nm). There was no effect from the Parafilm, and there was also found to be negligible signal from an empty petri dish container, i.e. in the absence of sample.

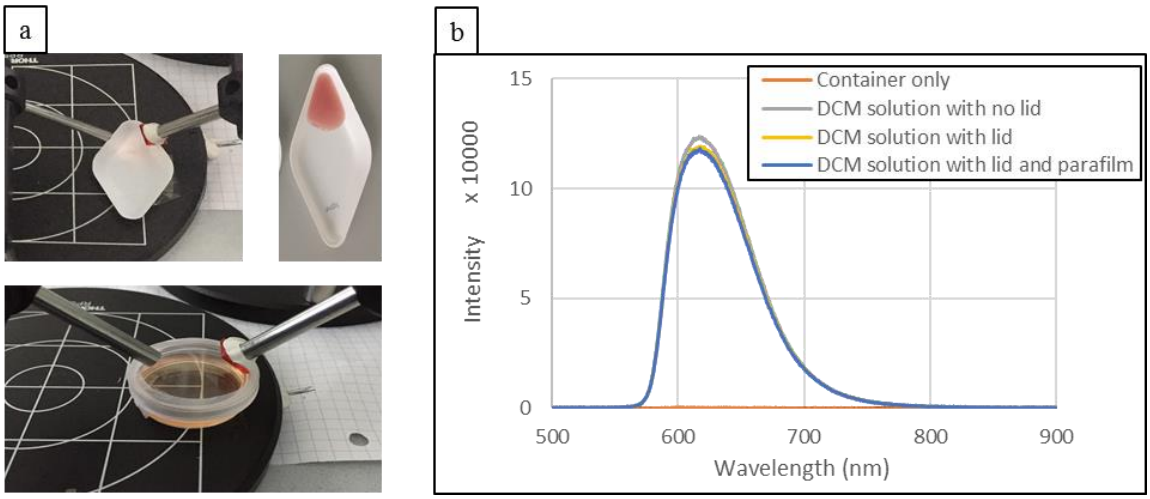


Figure 5.5 - Fluorescence experimental setup. a) Containers used for solutions: measuring boat with no lid (upper images) and micro petri dish with lid (lower image. b) Effect of lid on spectra intensity.

5.4.2 PpIX time series and photochemical degradation

A solution of porphyrin IX powder (Enzo Life Scientific, USA) was made as per manufactures guidelines[268]. 1 mg PpIX was dissolved into 5 mL 0.1 M sodium hydroxide. An equal volume (5 mL) of ethanol was then added giving a stock concentration of 0.1 mg/mL. The solution was stirred for 5 minutes until all the powder was dissolved. This stock solution was then diluted as required using distilled water. The solutions were made up on the day of use and were wrapped in foil and stored in a 5° refrigerator to limit any photochemical reaction until use.

The aim of the time series experiment was to assess if there was a change in the PpIX fluorescence spectrum as a function of exposure time.

The optical setup outlined above was used with the Thorlabs LED light source set at maximum intensity level. A x10 dilution of the stock PpIX solution was placed in the measuring boat. In addition, a "control" solution was made containing just the solvents (5 mL sodium hydroxide and 5 mL ethanol) in a x10 dilution using distilled water. This control solution also underwent the time series experiment to see if there was any change in background fluorescence from components other than the PpIX. A spectrum was measured immediately after the sample was illuminated (time = 0 s) and then a spectrum was collected every 30 seconds for 21600 seconds (6 hours). The exposure time was 1 s per spectrum.

Figure 5.6 shows the change in PpIX fluorescence over time. There is a dramatic change in both intensity and spectral shape over time. The change starts to occur at very short times, with a decrease in intensity and shift in the peak wavelengths evident after 30 seconds of exposure (Figure 5.6b). Further time series experiments, not shown here, which measured spectra every 1 second showed spectral change immediately after 1 to 2 seconds.

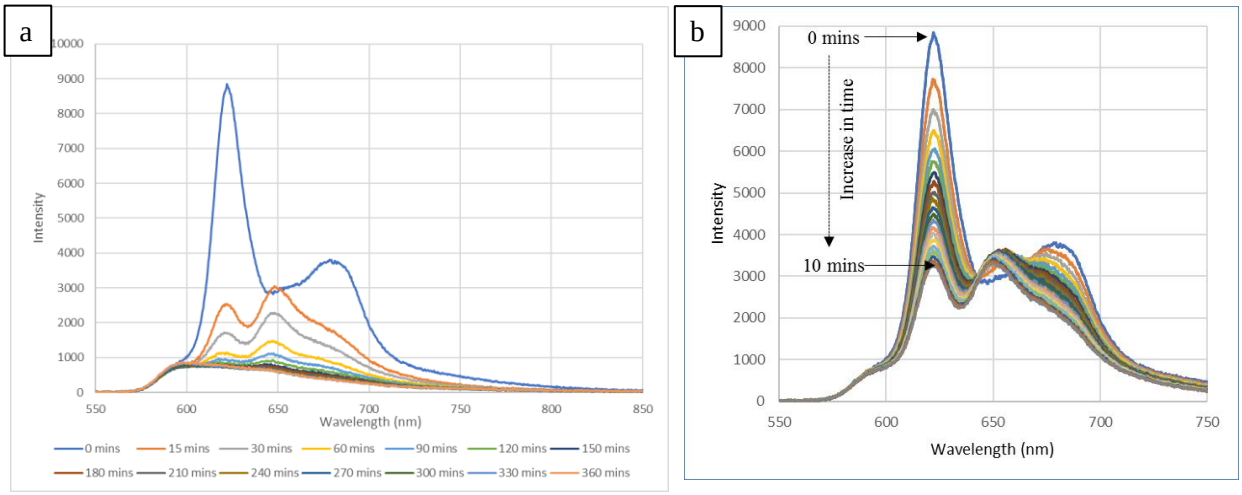


Figure 5.6- PpIX fluorescence time series. a) Spectra every 15 minutes from 0 to 60 minutes and then every 30 minutes thereafter up to 360 minutes (6 hours). b) Spectra every 30 s from 0 to 10 minutes.

A curve fitting model was built for the time series (Figure 5.7). Reference spectra are shown in Figure 5.7a, and consist of a background spectrum of the solvent (measured using the "control" solution), a PpIX spectrum (the first spectrum of the time series with the solvent spectrum subtracted) and the product spectrum (derived from the last spectra in the time series with the solvent and the reactant contribution subtracted). The curve fit modelling confirms that there appears to be a rapid photochemical reaction that occurs as soon as the PpIX solution is illuminated. The spectral contribution from the reaction product outweighs the contribution from the PpIX after just 2 minutes exposure time (Figure 5.7 (inset)).

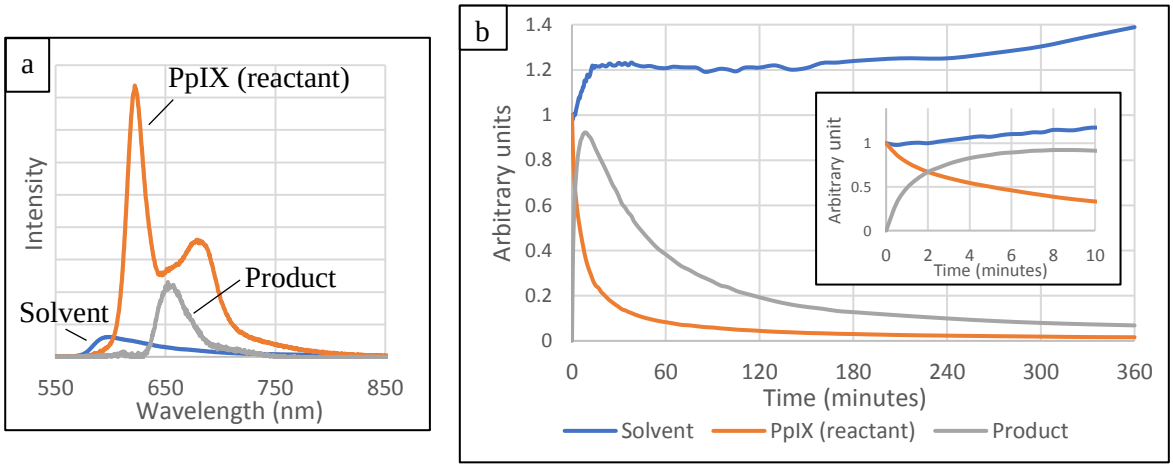


Figure 5.7 - Curve fit modelling of PpIX time series. a) Reference spectra used for curve fit model. b) Curve fit model of time series over 360 minutes. (inset: curve fit over first 10 minutes)

The reaction observed in the time series represents a light-induced, or photochemical, reaction of the PpIX. This degradation of PpIX during irradiation, into photoproducts absorbing and fluorescing at slightly longer wavelengths, has been observed in solution[269-271] and in cell cultures incubated with 5ALA[272-275]. In fact, the photochemical reaction has been exploited clinically as the basis for using

5ALA-induced PpIX fluorescence in photodynamic therapy[127,276,277] with the photochemical reaction products having a cytotoxic effect on cells with high accumulation of 5ALA. The predominant reaction

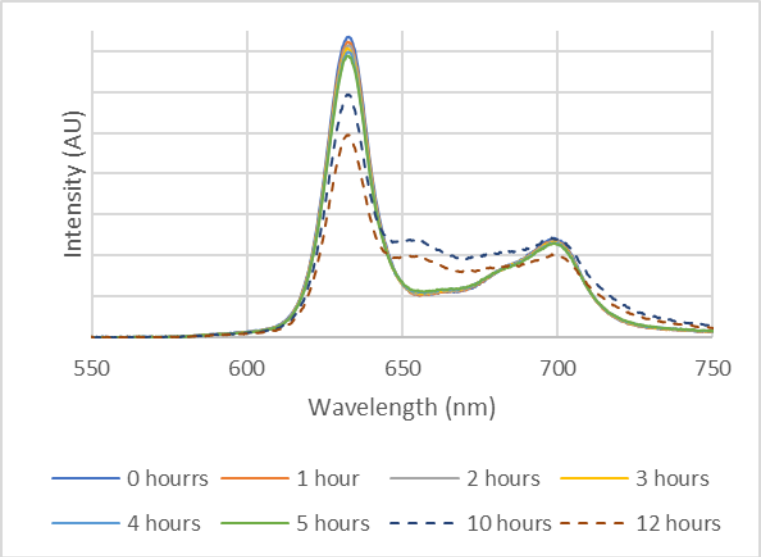


Figure 5.8 - PpIX time series using petri dish with lid. AU=arbitrary units

is an oxidation producing reactive oxygen species such as hydrogen peroxide and hydroxyl radicals[127,278-280]. Removing the oxygen from the PpIX solution has been shown previously to prevent the reaction[280,281]. When the time series experiment above was repeated using the petri dish with lid, the speed of the photochemical reaction was substantially attenuated (Figure 5.8). In this case, evidence of a photochemical reaction did not occur until approximately 10 hours of light exposure. The peak intensity of the PpIX spectra, namely the intensity peak at 630 nm, did decrease over a shorter time frame, losing nearly 5% intensity over 20 minutes. Although this significantly improves the stability of the solution, the observed decrease in intensity is still a confounding factor if using PpIX for laboratory-based fluorescence experiments where the fluorescence intensity is to be compared.

5.4.3 Summary

PpIX is very sensitive to light exposure, making its use for fluorescence intensity comparison experiments difficult. For such applications, a new stock solution of PpIX would need to be made up on the day of experimentation, and a different sample used for each fluorescence measurement. It would also be imperative to make sure that the spectra are recorded as soon as the excitation light illuminates the sample, as the degradation in the sample starts immediately. These precautions, whilst achievable (they are employed in further experiments using PpIX described later), are not generally practical, and introduce a substantial risk of error when measuring fluorescence intensities.

It was therefore decided that an alternative fluorescence standard should be considered that matched the absorption and fluorescence wavelengths of PpIX but that was photochemically stable, allowing for increased reliability in further fluorescence intensity experiments.

5.5 DCM laser dye

Laser dyes are stable organic compounds used in dye lasers. In a dye laser the excitation laser source, known as the pump laser, is passed through the laser dye solution, which absorbs the light and emits at a longer wavelength. This allows for the laser to emit various wavelengths depending on the laser dye used[282,283]. By their very nature laser dyes are photochemically stable and have constant absorption and emission profiles, making them ideal to act as a fluorescence standard in fluorescence intensity comparison experiments. For this study, the laser dye 4-Dicyanomethylene-2-methyl-6-p-dimethylaminostyryl-4H-pyran (DCM)[284,285] dissolved in ethanol was selected to best match the absorption and fluorescence wavelengths of PpIX.

A solution of DCM in ethanol was made with a concentration in the order of 10^{-5} mol L^{-1}, similar to that used for its laser dye application[286]. 1.517 mg of DCM powder (Photonic Solutions, UK) was dissolved in 500 mL ethanol giving a concentration of 1.00 x 10^{-5} mol L^{-1}[xx].

The DCM and PpIX absorption spectra (measured using transmission spectroscopy) and fluorescence profile (measured using the fluorescence experimental setup as outlined above) are shown in Figure 5.9. DCM absorption occurs in the 400 nm to 550 nm region, with maximum absorption at 460 nm, compared to PpIX in the 350 nm to 450 nm range with maximum absorption at 405 nm. Both will be excited by the Thorlabs LED light source and the operative light sources (emission range depicted by shaded area in Figure 5.9a). With regards to fluorescence, DCM has a maximum fluorescence peak at 620 nm, very similar to the maximum PpIX fluorescence peak at 635 nm (Figure 5.9b).

A time series experiment was also carried out on the DCM solution using the same experimental setup as described for the PpIX time series and using the petri dish with a lid sealed with Parafilm. There was

[xx] Calculated using: **Mass (g) = Molarity *(moles/L)* x Molar mass *(g/mole)* x volume *(L)***. Therefore **Molarity = Mass / (Molar mass x Volume)** with Molar mass DCM = 303.37 g/mole
Molarity = 0.001517 / (303.37 x 0.5) = 1.00 x 10^{-5}

less than 1% change in the DCM fluorescence intensity over the 12-hour experimental period, confirming the photochemical stability of this compound (Figure 5.10).

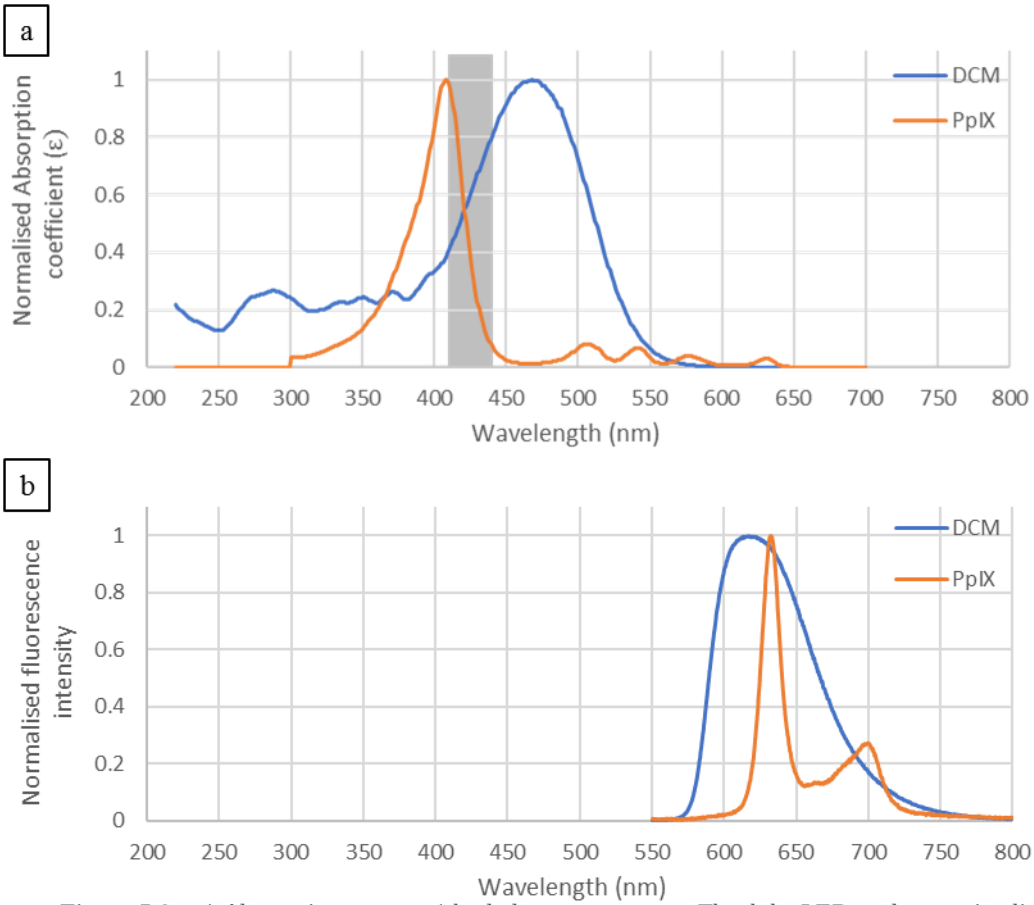


Figure 5.9 - a) Absorption spectra (shaded area represents Thorlabs LED and operative light sources wavelength range) and b) fluorescence spectra for DCM (in ethanol) and PpIX (in NaOH and ethanol)

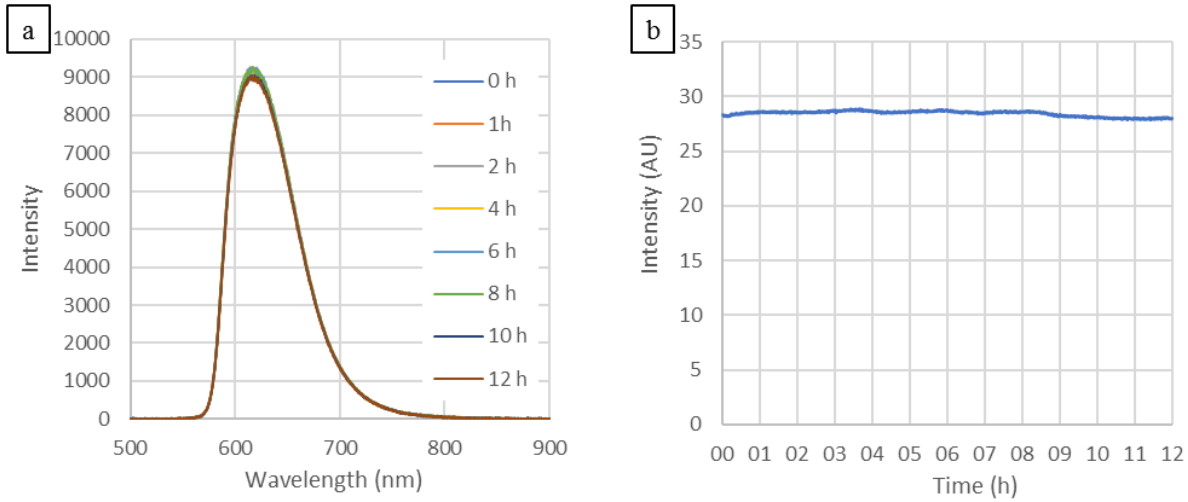


Figure 5.10 - DCM time series. a) Fluorescence spectra over 12 hours. b) Change in total fluorescence intensity over time

DCM provides a stable fluorescence standard that can be used for fluorescence intensity experiments as a surrogate for PpIX fluorescence. For the DCM fluorescence measurements to be meaningful it must be possible to correlate the DCM fluorescence with the equivalent PpIX fluorescence for known concentrations of each substance. Prior to deriving this correlation experimentally, however, it is important to establish that there is no signal saturation for the concentrations and light intensities used, which would compromise the measured correlation. To this end, experiments to look for evidence of signal saturation are described in the next section prior to reporting the method and results of the correlation of DCM and PpIX fluorescence.

5.6 Investigating evidence of signal saturation

The fluorescence signal from a solution should be linearly proportional to both the light intensity and the sample concentration (Figure 5.11). There are certain conditions under which this is no longer true, and the signal is said to be 'saturated'. The first of these is when, at a fixed light intensity, all the photons from the light source are absorbed by the sample. Under these conditions, increasing the concentration

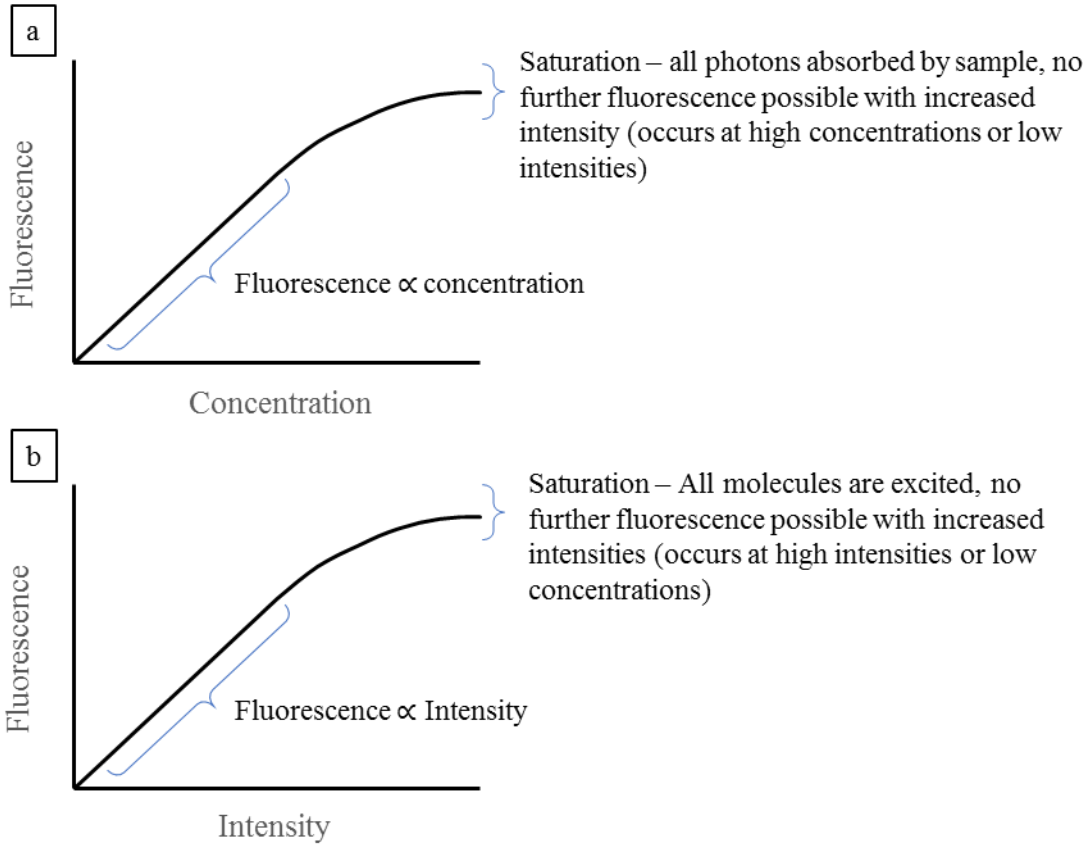


Figure 5.11 – Relationship between a) fluorescence and concentration and b) fluorescence and light intensity

of the sample will not lead to any more photons being absorbed, and therefore will not lead to an increase in fluorescence. This condition can occur at high sample concentrations and/or low light intensities (Figure 5.11a). The second situation in which saturation can occur is when, at a fixed sample concentration, all the molecules in the sample are excited by photons from the light source. In this case, increasing the light intensity further will not lead to an increase in fluorescence, as there are no more ground-state molecules available to absorb the additional photons. This condition can occur at high light intensities and/or low sample concentrations (Figure 5.11b).

In any experiment that attempts to relate fluorescence intensity to sample concentration, it is important to make sure that there is no signal saturation for the concentrations and light intensities used. Fluorescence spectra of PpIX and DCM solutions were therefore recorded using different solution concentrations and different light intensities to establish if there was a linear relationship between these variables or if there was evidence of fluorescence saturation. The results of these experiments were then used to decide the optimum concentrations and light intensities to be used for the PpIX and DCM fluorescence correlation experiments.

5.6.1 Method

The fluorescence experimental setup outlined above was used along with the Thorlabs LED light source at 12 mm working distance. The light source intensity was changed according to the arbitrary scale on the light source dial from 1 to 6 which corresponded to a non-linear increase in intensity as illustrated in Figure 5.3c. At each light intensity a fluorescence spectrum was recorded, yielding six fluorescence spectra per sample (see Figure 5.12 and Figure 5.13 for examples). The total fluorescence intensity was calculated for each spectrum and plotted against the corresponding light intensity (V) to give a fluorescence versus intensity plot, similar to that shown schematically in Figure 5.11b, for each concentration of DCM and PpIX.

The DCM and PpIX stock solutions were made using the methods outlined in sections 5.5 and 5.4.2 respectively. Several DCM and PpIX stock solutions were made and used in this and/or later experiments. Table 5.3 summarises each solution name and corresponding concentration.

Solution name	Mass of solute (g)	Volume of solvent (dm^3)	Molarity $(mole\ dm^{-3}) = \frac{Mass}{(Molar\ mass \times volume)}$
DCM (molar mass = 303.37 g/mol)			
DCM solution 1	0.002990	0.2	4.93×10^{-5}
DCM solution 2	0.003150	0.2	5.19×10^{-5}
DCM previous solution*	0.001517	0.5	1.00×10^{-5}
PpIX (molar mass = 562.67 g/mol)			
PpIX solution 1	0.005590	0.2	4.97×10^{-5}
PpIX solution 2	0.005810	0.2	5.16×10^{-5}

*Table 5.3 - DCM and PpIX stock solutions. The composition of each stock solution is shown with the calculated concentration (molarity). *DCM previous solution was a solution of DCM that has been made up 4 weeks prior to its use in the experiments presented here.*

From these stock solutions several serial dilutions were performed, and selected concentrations used in this and/or later experiments. The choice of dilution factor and the number of dilutions were dictated by experiments reported later in this chapter which investigate thresholds of fluorescence detection by operative light sources. For DCM, x2, x10 and x1.2 serial dilutions were used and x2, x10 and x1.5 serial dilutions were used for PpIX. In each dilution series every dilution created was given a label and a corresponding relative concentration as a percentage of the stock solution (stock solution = 100% relative concentration). For clarity, these labels and relative concentrations are used from here on rather than the actual concentrations. Table 5.4 summarises the DCM and PpIX solutions created from the various dilution series. The true concentrations can be calculated using the stock solution molarity (from Table 5.3) multiplied by the percent relative concentration (from Table 5.4).

DCM dilution series Stock solution: DCM 1 (100% concentration)		
Label	% Conc.	Dilution series
DCM 1a	50.000	x2 from stock
DCM 1b	25.000	x2
DCM 2	10.000	x10 from stock
DCM B	8.330	x1.2 from stock
DCM C	6.940	x1.2
DCM D	5.790	x1.2
DCM E	4.820	x1.2
DCM F	4.020	x1.2
DCM G	3.350	x1.2
DCM H	2.790	x1.2
DCM I	2.330	x1.2
DCM J	1.940	x1.2
DCM K	1.620	x1.2
DCM L	1.350	x1.2
DCM M	1.120	x1.2
DCM 3	1.000	x10
DCM N	0.935	x1.2
DCM O	0.779	x1.2
DCM P	0.649	x1.2
DCM Q	0.541	x1.2
DCM R	0.451	x1.2
DCM S	0.376	x1.2
DCM T	0.313	x1.2
DCM U	0.261	x1.2
DCM V	0.217	x1.2
DCM W	0.181	x1.2
DCM X	0.151	x1.2
DCM Y	0.126	x1.2
DCM Z	0.105	x1.2
DCM 4	0.100	x10

PpIX dilution series Stock solution: PpIX 1 (100% concentration)		
Label	% Conc.	Dilution series
PpIX 1a	50.000	x2 from stock
PpIX 1b	25.000	x2
PpIX 2	10.000	x10 from stock
PpIX A	6.670	x1.5 from stock
PpIX B	4.440	x1.5
PpIX C	2.960	x1.5
PpIX D	1.980	x1.5
PpIX E	1.320	x1.5
PpIX 3	1.000	x10
PpIX F	0.878	x1.5
PpIX G	0.585	x1.5
PpIX H	0.390	x1.5
PpIX I	0.260	x1.5
PpIX J	0.173	x1.5
PpIX K	0.116	x1.5
PpIX 4	0.100	x10

Table 5.4 - Serial dilutions of DCM and PpIX showing solution label and the percent concentration (% Conc.) relative to stock solution. The dilution series used to derive that concentration is also shown with the first dilution in that series relative to the stock solution and then subsequent dilutions relative to the previous row from the same series.

In these saturation experiments the following stock solutions used were: DCM solution 1, DCM solution 2, PpIX solution1 and PpIX solution 2. The DCM solutions used (and relative concentrations) were: DCM 1 (100%), DCM 1a (50%), DCM 2 (10%), DCM E (4.820%), DCM I (2.330%), DCM M (1.120%), DCM 3 (1.000%) and DCM N (0.935%). The PpIX solutions used (and relative concentrations) were: PpIX 1 (100%), PpIX 1a (50%), PpIX 1b (25%), PpIX 2 (10%), PpIX A (6.670%), PpIX B (4.440%),

PpIX C (2.960%), PpIX D (1.980%), PpIX E (1.320%), PpIX 3 (1.000%), PpIX F (0.878%), PpIX G (0.585%), PpIX H (0.390%), PpIX I (0.260%), PpIX J (0.173%) and PpIX 4 (0.100%).

For each solution the total fluorescence intensity was calculated and plotted against the corresponding relative concentration (%) to give a fluorescence versus concentration plot, similar to that shown schematically in Figure 5.11a, for each intensity setting.

5.6.2 Results

Similar results were observed using all DCM and PpIX stock solutions. The results from DCM solution 1 and PpIX solution 1 are reported here.

There was no evidence of DCM saturation with increasing light intensity. Figure 5.12 shows fluorescence spectra at each light intensity (Figure 5.12a, c and e) and the calculated total fluorescence intensity for each spectrum plotted against the light intensity (Figure 5.12b, d and f). Only three concentrations of DCM, across the concentration range used, are shown DCM 1, DCM I and DCM N. All other concentrations showed similar results. As would be expected, there was a greater contribution from background noise at lower concentrations when the overall fluorescence signal is weak (Figure 5.12e).

Similar results were seen with PpIX: a linear relationship was observed between fluorescence and light source intensity, with no evidence of saturation across a range of concentrations. This is illustrated in Figure 5.13 which shows fluorescence spectra and total fluorescence versus light intensity plots for three concentrations of PpIX. The lowest concentration shown here is PpIX G (0.585% relative concentration). At concentrations lower than this the PpIX signal could not be extracted from the noise and therefore these concentrations were excluded. A linear relationship was seen in all solution concentrations measured above PpIX G.

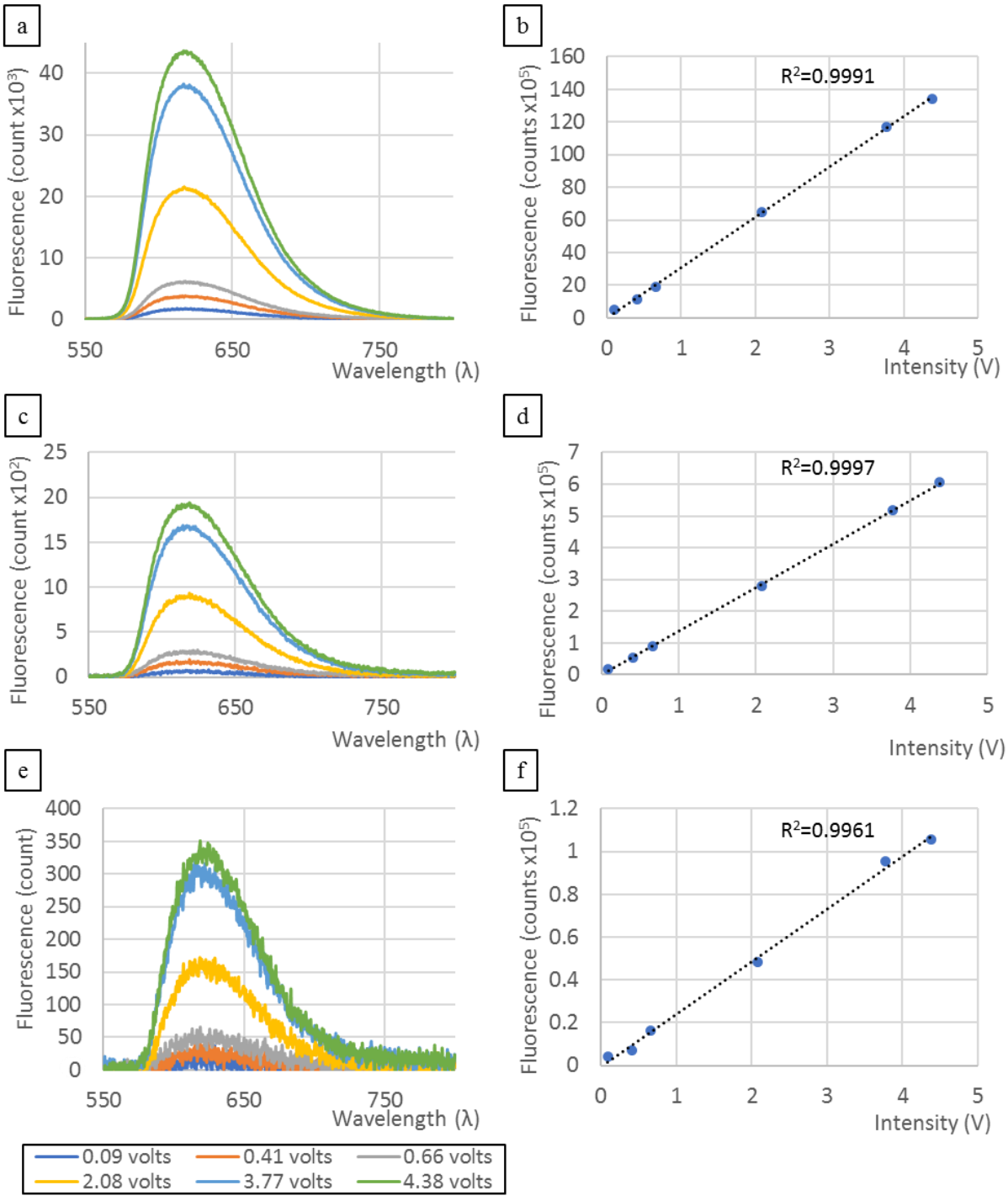


Figure 5.12 – Fluorescence versus intensity for DCM. Fluorescence spectra at each light intensity (left) and corresponding total fluorescence versus light intensity (right) for given concentrations of DCM. Three concentrations shown: DCM 1 [a) and b)], DCM I [c) and d)] and DCM N [e) and f)]. Dotted line represents fitted linear regression model with corresponding R^2 value.

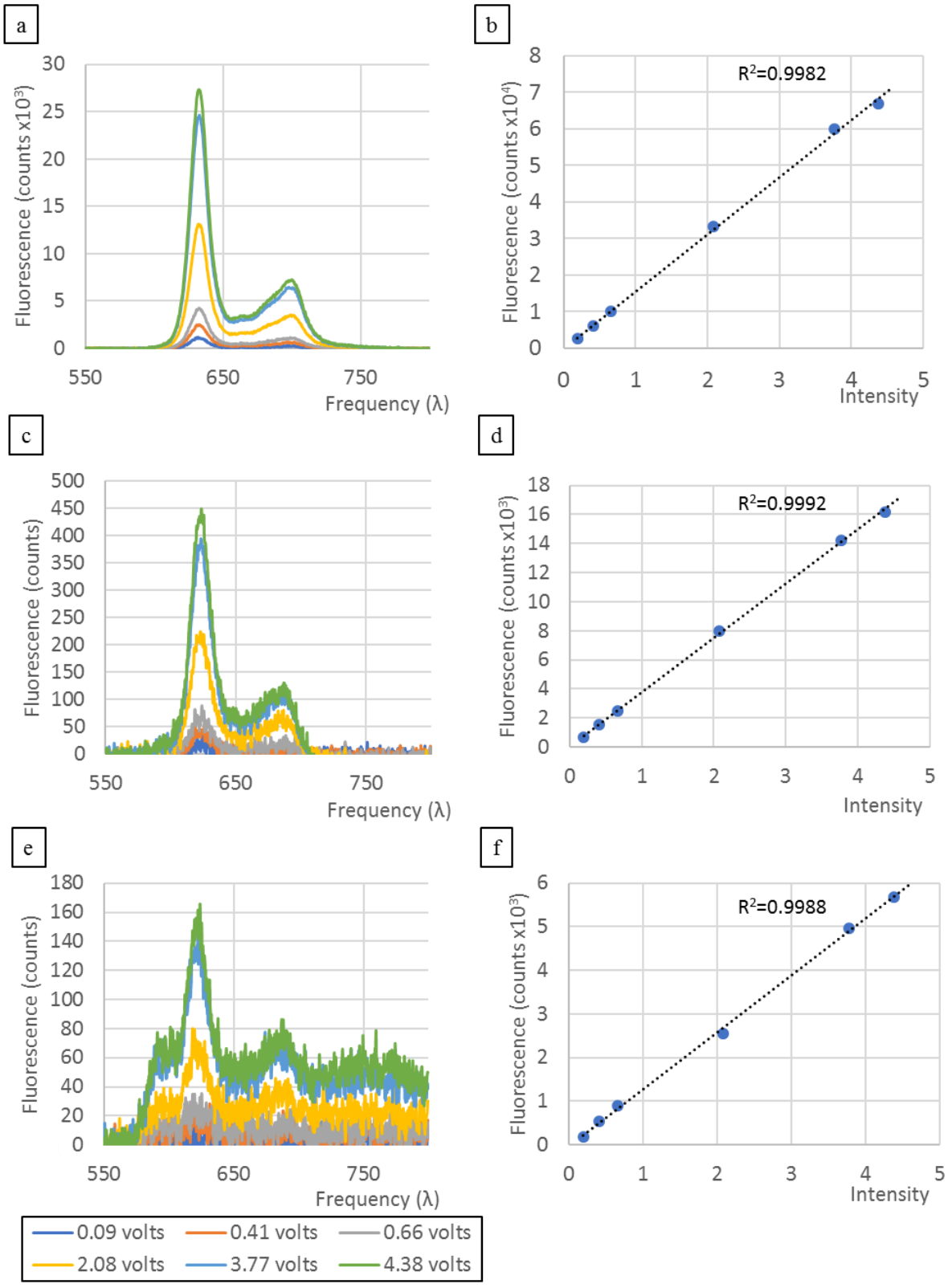


Figure 5.13 - Fluorescence versus intensity for PpIX. Fluorescence spectra at each light intensity (left) and corresponding total fluorescence versus light intensity (right) for given concentrations of PpIX. Three concentrations shown: PpIX 1 [a) and b)], PpIX B [c) and d)] and PpIX G [e) and f)]. Dotted line represents fitted linear regression model with corresponding R^2 value.

The total fluorescence was also plotted against solution concentration for each light intensity setting. Figure 5.14 and Figure 5.15 depict the resulting plots for three light intensity settings – 4.38 V (maximum light intensity), 2.08 V and 0.09 V (minium light intensity) – for solutions of DCM and PpIX respectively. The fluorescence intensity is also plotted against the $\log_{10}$ of the relative concentration to better visualise the relationship at lower concentrations. At higher concentrations, DCM fluorescence appears to be linearly related to the concentration at all light intensities (Figure 5.14a, c and e). In the very low concentration range there appears to be a fall away from the linear relationship, with fluorescence lower than predicted (most easily seen in the $\log_{10}$ plots: Figure 5.14b, d and f). This is not characteristic of a saturation effect but rather an effect of reaching the signal-to-noise limit of the experimental setup; the spectrometer is unable to distinguish the fluorescence signal at these low concentrations over the background. In this concentration range it is not possible to definitively determine if the relationship of fluorescence with concentration is linear with the current experimental setup, and therefore these low concentrations were not used in DCM calibration experiments reported later. This effect is not seen in the PpIX results (Figure 5.15). Here, however, there is some evidence of fluorescence saturation at the 100% relative concentration, with the total fluorescence intensity plateauing at this concentration. This would suggest that over this high concentration range all of the incident photons are completely absorbed by the solution, and a further increase in concentration does not result in a proportional increase in fluorescence (see Figure 5.11a). The plateauing becomes more apparent as the light intensity falls (compare Figure 5.15a, c and e). This supports the theory that this is a saturation effect with the photons from the lower light intensity being completely absorbed at lower concentrations to those from the higher light intensity. Despite this possible saturation effect at higher concentrations, a linear regression model still shows an excellent goodness of fit with a very high R^2 value. It might therefore be considered overly cautious to avoid these high concentrations of PpIX in later calibration experiments. Nonetheless, to eliminate this as a confounding factor, these high concentrations were avoided in calibration experiments reported later.

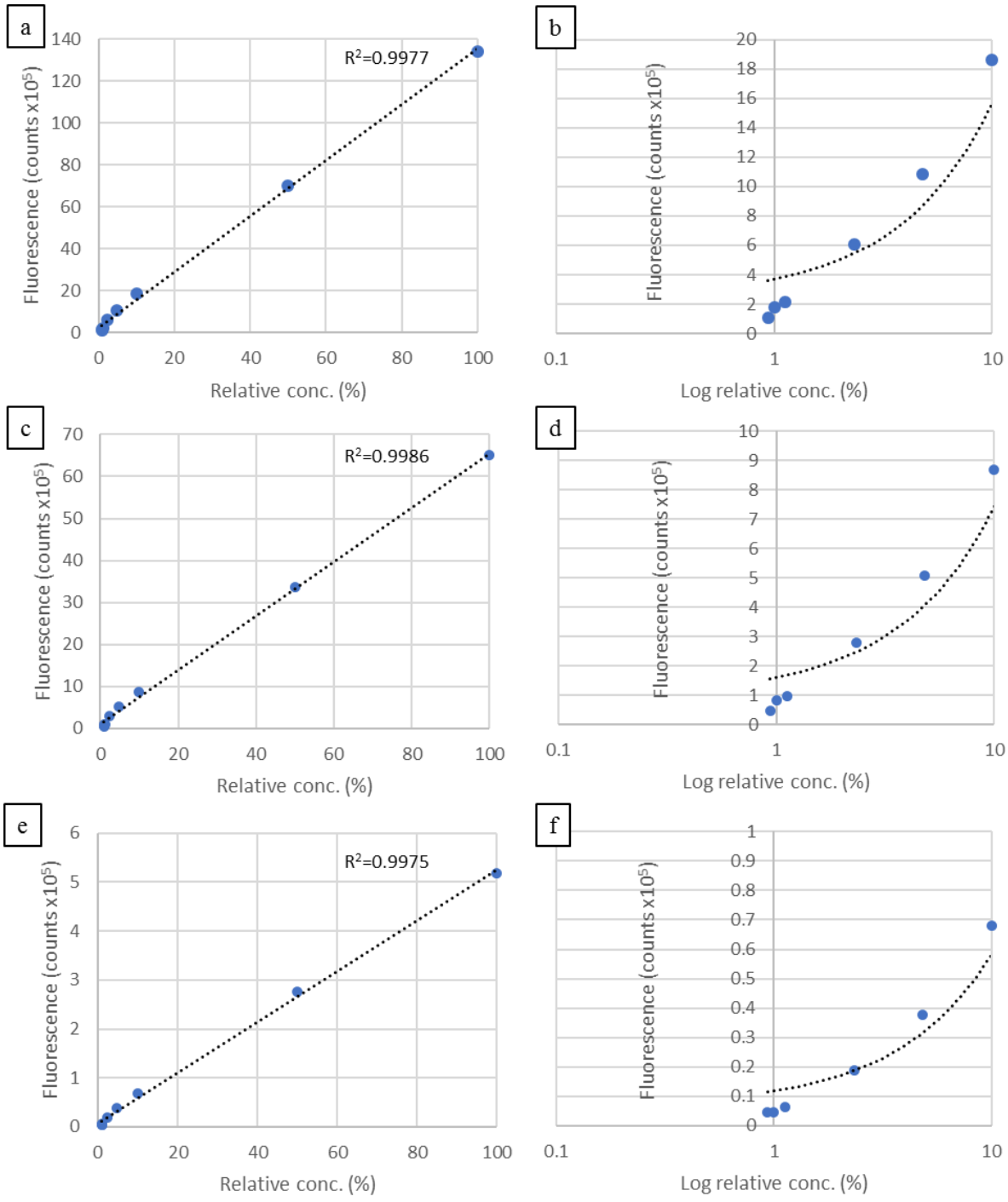


Figure 5.14 – Fluorescence versus concentration for DCM. Total fluorescence intensity versus relative concentration (left) and $\log_{10}$ of relative concentrations (right; 100% and 50% relative concentration omitted for clarity) for three different light intensity settings: 4.38 V [a) and b)], 2.08 V [c) and d)] and 0.09 V [e) and f)]. Dotted line represents fitted linear regression model with corresponding R^2 value.

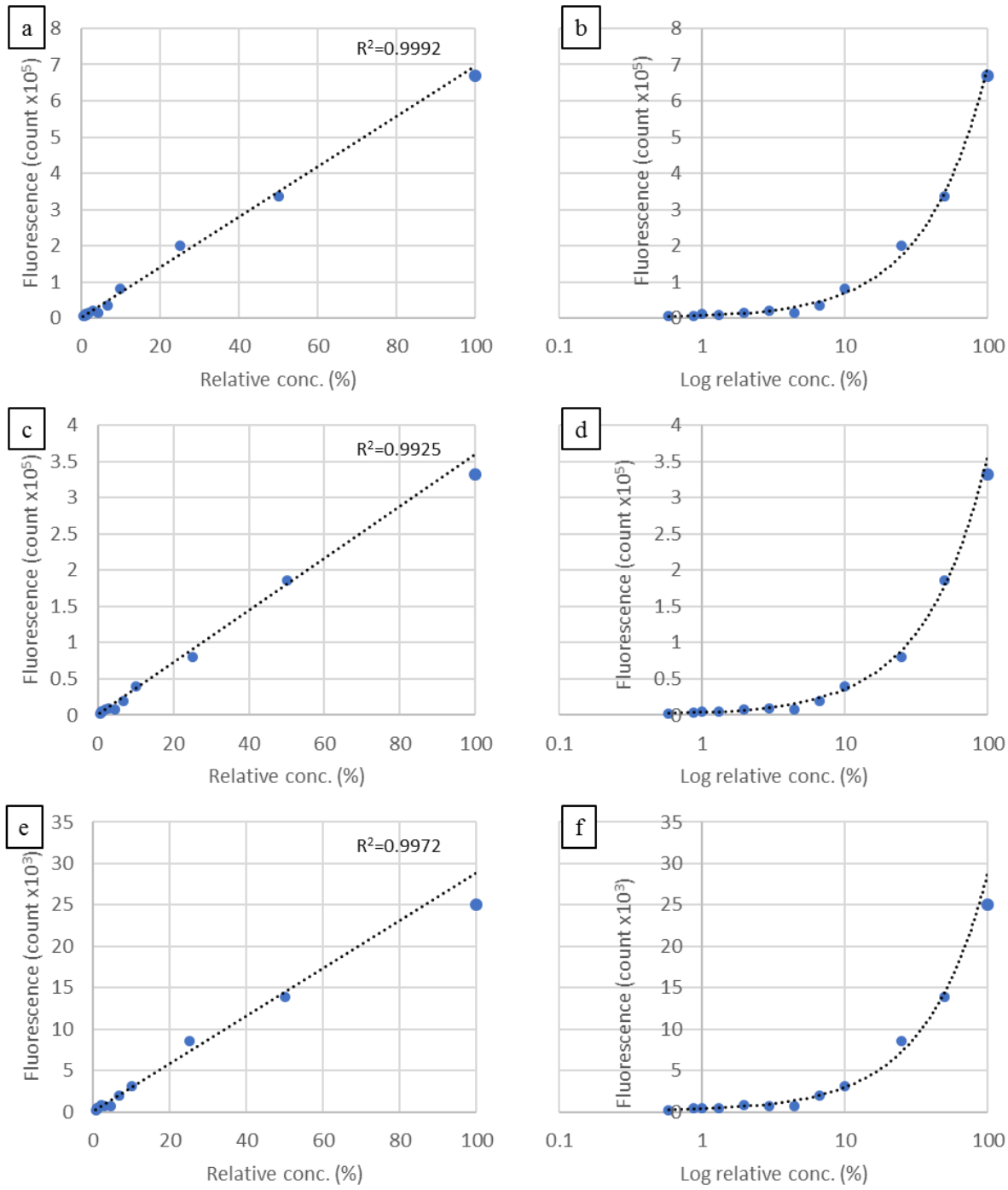


Figure 5.15 - Fluorescence versus concentration for PpIX. Total fluorescence intensity versus relative concentration (left) and $\log_{10}$ of relative concentrations (right) for three different light intensity settings: 4.38 V [a) and b)], 2.08 V [c) and d)] and 0.09 V [e) and f)]. Dotted line represents fitted linear regression model with corresponding R^2 value.

5.6.3 Summary

These data confirm that there is no evidence of fluorescence saturation with increasing light intensity in either DCM or PpIX. There is however some evidence that the use of very low concentrations of DCM or very high concentrations of PpIX would be working outside the range where fluorescence intensity is linearly proportional to solution concentration. The experimental correlation of DCM with PpIX fluorescence, reported next, therefore avoided the use of concentrations within these ranges.

5.7 Correlation of PpIX and DCM fluorescence

As previously stated, to use DCM as a surrogate for PpIX in fluorescence experiments it must be possible to determine the concentration of PpIX that would give an equal fluorescence intensity to that from a known concentration of DCM. This requires the determination of the ratio between DCM fluorescence intensity (F_1) and PpIX fluorescence intensity (F_2) as a function of the two concentrations. This ratio, referred to henceforth as the fluorescence ratio ($\frac{F_1}{F_2}$), will be different for given concentrations of each solution and for each light source used (Thorlabs LED, microscope or endoscope).

We first use the Beer-Lambert law to derive an expression for the fluorescence ratio. The predictions of the model will then be compared with fluorescence ratios derived from experimental measurements of fluorescence of different concentrations of DCM and PpIX using the Thorlabs LED, endoscope and microscope light sources. A close correlation should be found between the theoretically derived ratios and the experimental measurements of fluorescence.

5.7.1 Derivation of fluorescence intensity ratio

The fluorescence intensity ratio ($\frac{F_1}{F_2}$) for DCM (F_1= fluorescence intensity of DCM) and PpIX (F_2= fluorescence intensity of PpIX) can be expressed as:

$$\frac{F_1}{F_2} = \frac{I_{abs,1}\,\Phi_1}{I_{abs,2}\,\Phi_2} \qquad [6]$$

where I_{abs} = intensity of absorbed photons and Φ_1 and Φ_2 are the fluorescence quantum yields for DCM and PpIX, respectively. The quantum yield is defined as the number of times a specific event, in this case fluorescence, occurs per photon absorbed by the substance. The fluorescence quantum yield

can therefore be defined as the ratio of the number of photons emitted to the number of photons absorbed. The fluorescence quantum yields for DCM (Φ_1) and PpIX (Φ_2) are 0.43[287] and 0.06[288] respectively.

From the Beer-Lambert law I_{abs} is defined as

$$I_{abs} = I_0 - I \quad [7]$$

$$= I_0\left(1 - 10^{-\varepsilon c l}\right) \quad [8]$$

where I = transmitted intensity, I_0 = incident intensity, c = concentration of solution (mol dm^{-3}), l = length of solution the light passes through (cm) and ε = molar extinction coefficient (dm^3 mol^{-1} cm^{-1}). The molar extinction coefficient (also known as the absorption coefficient) is a measure of how well the molecule of interest absorbs at a given wavelength of light. Transmission spectra of DCM and PpIX solutions of known concentrations ($c_1$and c_2 respectively) can be measured using standard transmission spectroscopy. Values of ε_1 and ε_2 can then be calculated from the percentage of light transmitted (T) at a given wavelength (λ) using a rearrangement of equation [8] giving:

$$\varepsilon(\lambda) = \frac{-\log\frac{T(\lambda)}{100}}{c\,l} \quad [9]$$

Combining equations [6] and [8], for a single fixed wavelength and illumination intensity, gives:

$$\frac{F_1}{F_2} = \frac{\left(1 - 10^{-\varepsilon_1(\lambda)\,c_1 l_1}\right)\Phi_1}{\left(1 - 10^{-\varepsilon_2(\lambda)\,c_2 l_2}\right)\Phi_2} \quad [10]$$

For broadband light source excitation, I_0, ε_1 and ε_2 become functions of wavelength and thus we need to integrate over the spectrum of the incident light. If we assume that Φ_1 and Φ_2 are independent of the wavelength[xxi] then, for a broadband light source:

$$\frac{F_1}{F_2} = \frac{\Phi_1 \int I_0(\lambda)\left(1 - 10^{-\varepsilon_1(\lambda)\,c_1 l_1}\right)d\lambda}{\Phi_2 \int I_0(\lambda)\left(1 - 10^{-\varepsilon_2(\lambda)\,c_2 l_2}\right)d\lambda} \quad [11]$$

For each molecule, values of $\varepsilon_1(\lambda)$ and $\varepsilon_2(\lambda)$ can be measured experimentally as a function of wavelength (see Section 5.7.2). With these values, it is now possible to solve equation [11], for given concentrations of DCM and PpIX ($c_1$and c_2 respectively), in order to calculate the fluorescence ratio ($\frac{F_1}{F_2}$). Similarly, if we measure fluorescence from a known concentration of DCM (c_1) we can calculate the concentration

[xxi] The fluorescence quantum yield, Φ, is normally independent of wavelength since fluorescence is usually from the lowest excited electronic state following rapid relaxation from higher-lying states.

of PpIX (c_2) that would give the same intensity of fluorescence as, in this case, the fluorescence ratio would be equal to one ($\frac{F_1}{F_2} = 1$). The only unknown is c_2 and this can therefore be calculated from equation [11] using a non-linear fit model (the Microsoft Excel add-in, 'Solver'[266] was used here).

5.7.2 Method

In order to use equation [11], the two extinction coefficients $\varepsilon_1(\lambda)$ and $\varepsilon_2(\lambda)$ must be known as a function of wavelength. These quantities are straightforward to measure using a standard laboratory UV-vis spectrometer. The measurements were made on stock solutions DCM solution 1, DCM solution 2, DCM solution previous, PpIX solution 1 and PpIX solution 2. The data collection and analysis were carried out as follows:

1. Each solution of DCM and PpIX of known concentration ($c_1$and $c_2$respectively) was placed in a 1 cm path length cuvette, and the percentage transmission across the wavelength range of 290 nm to 750 nm was measured using a UV-Vis spectrometer (Thermo Scientific, UK). To prevent spectrometer saturation, both DCM and PpIX solutions required dilution of the stock solution (x5 dilution and x10 dilution, respectively). Two spectra were collected for each solution and a mean taken.
2. Values of ε were calculated at each wavelength from the mean percent transmission spectrum (T) of each solution using equation [9] above (l = 1 cm).
3. The mean values of ε at each wavelength were calculated from all solutions of DCM and PpIX to give a mean ε spectrum per substance per wavelength.
4. Light source spectra (normalised to area under curve = 1) from the endoscope, microscope and Thorlabs LED were each used to calculate $\frac{F_1}{F_2}$ values integrated across the whole wavelength range. These were then divided by the integral of the light source spectra (equation [11]) to give an overall value of $\frac{F_1}{F_2}$ for each light source at known values of c_1 and c_2. This calculation was applied using the concentrations of DCM and PpIX used in the experimental measurements to generate a derived value of $\frac{F_1}{F_2}$ which could then be compared with the experimental value.

With regards to the experimental measurements of fluorescence spectra, all three light sources were set up as described in the light source calibration experiment and general methods. A working distance of 20 mm for the Thorlabs LED, 30 mm for the endoscope and 200 mm for the microscope were used. The maximum light intensity setting was used in each case.

Given the results of the saturation experiments reported in Section 5.6.2, low concentrations of PpIX were used compared to high concentrations of DCM. Different stock solutions were also used to show reproducibility of results. The combinations of stock solutions and concentrations of DCM and PpIX used are shown in Table 5.5. For each concentration of solution five spectra were taken and the mean spectrum calculated. The total intensity over the complete wavelength range was summed for each concentration of DCM and PpIX to give values of F_1 and F_2, thus allowing for an experimental estimate of the fluorescence intensity ratio ($\frac{F_1}{F_2}$) for that combination of solution concentrations. These ratios were then compared with the corresponding derived values for that concentration combination as outlined above.

5.7.3 Results

The mean extinction coefficients as a function of wavelength for all stock solution samples of DCM and PpIX, as derived from the transmission spectra, are shown in Figure 5.16. The experimental uncertainty in the raw transmission spectra, and thus the derived values of ε, is represented by one standard deviation in Figure 5.16. The uncertainty is greater for PpIX, which most likely reflects the photochemical instability of this solution compared to DCM, despite the practical measures taken to minimise light exposure of the solutions prior to transmission spectroscopy. The derived fluorescence ratios calculated using the measured extinction coefficients are shown for each light source in Table 5.5. The experimental values of F_1 and F_2 were calculated from the measured fluorescence spectra and are also shown in Table 6.5 for each combination of concentrations. The final column of the table shows the ratio of the experimentally measured and derived values as a measure of the correlation between experimental and derived values. The mean ratio was 1.03 (standard deviation $\sigma = 0.08$) using the Thorlabs LED, 0.65 ($\sigma = 0.16$) using the microscope, 1.04 ($\sigma = 0.12$) using the endoscope and 0.93 ($\sigma = 0.21$) overall. Although there is less correlation using the microscope compared to the other light sources, overall there is good

agreement between the experimental and derived values, and there was no significant difference between them (p=0.725, Wilcoxon rank sum test[xxii]). The use of "DCM previous solution" stock solution had comparable results to the other solutions, which would suggest there is no change in transmission characteristics if the DCM solution is stored. This supports the results of the DCM times series experiments, which confirmed DCM to be photochemically inert.

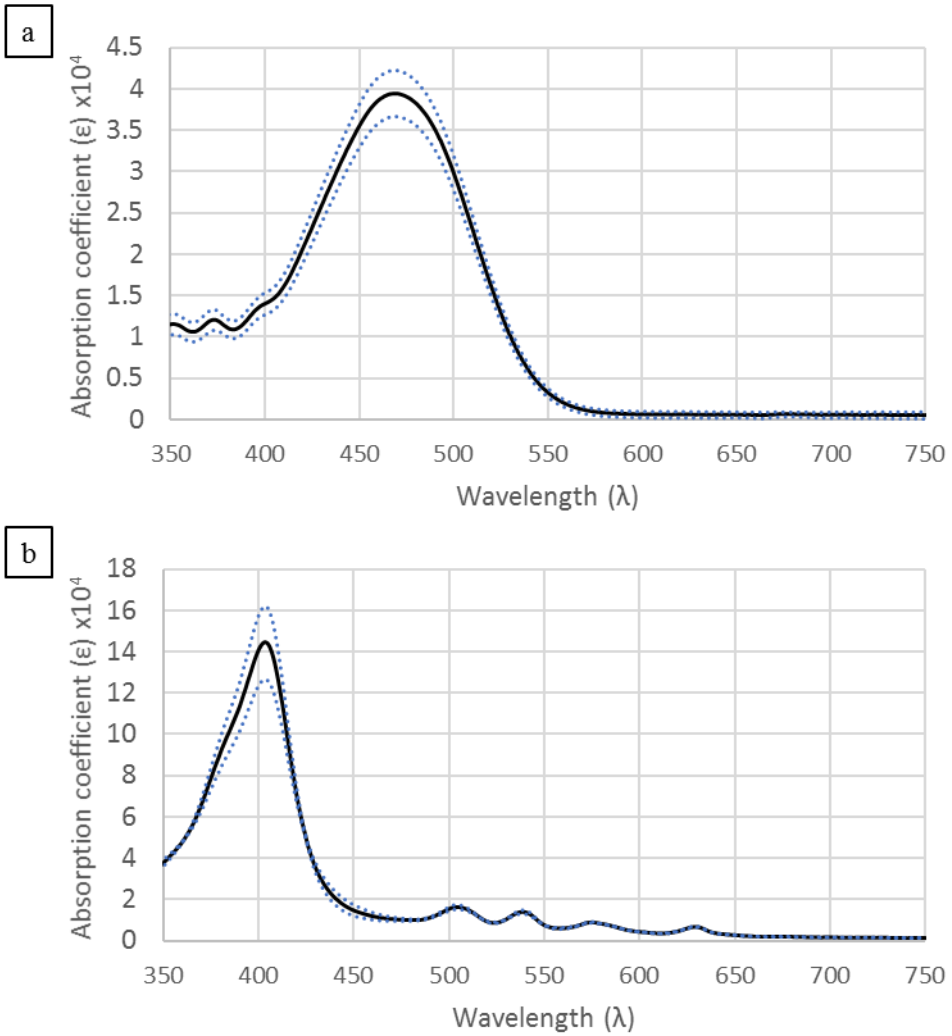


Figure 5.16 – Mean extinction coefficients. Mean values of extinction coefficients for a) DCM and b) PpIX for each wavelength. Dotted lines represent +/- one standard deviation.

[xxii] Wilcoxon rank sum test used here because this is non-parametric data of independent variables

Thorlabs LED

PpIX solution		DCM solution		Fluorescence ratio $\frac{F_1}{F_2}$		Ratio comparison: Experimental / Derived
stock	dilution	stock	dilution	Experimental	Derived	
Sol. 1	PpIX 3 (1%)	Sol. 1	DCM 1 (100%)	89.92	85.81	1.05
			DCM 2 (10%)	21.28	20.78	1.02
		Sol. 2	DCM 1 (100%)	88.93	86.74	1.03
			DCM 2 (10%)	20.50	21.74	0.94
		Prev. Sol.	DCM 1 (100%)	38.11	37.34	1.02
Sol. 1	PpIX D (1.98%)	Sol. 1	DCM 1 (100%)	49.14	45.71	1.08
			DCM 2 (10%)	13.35	11.05	1.21
		Sol. 2	DCM 1 (100%)	44.87	46.20	0.97
			DCM 2 (10%)	12.02	11.56	1.04
		Prev. Sol.	DCM 1 (100%)	18.12	19.86	0.91

Microscope

PpIX solution		DCM solution		Fluorescence ratio $\frac{F_1}{F_2}$		Ratio comparison: Experimental / Derived
stock	dilution	stock	dilution	Experimental	Derived	
Sol. 1	PpIX 3 (1%)	Sol. 1	DCM 1 (100%)	81.66	93.51	0.87
			DCM 2 (10%)	9.90	23.42	0.42
			DCM B (8.33%)	11.62	19.83	0.59
			DCM E (4.82%)	7.35	12.02	0.61
			DCM I (2.33%)	3.52	6.00	0.59
			DCM M (1.12%)	2.39	2.94	0.81

Endoscope

PpIX solution		DCM solution		Fluorescence ratio $\frac{F_1}{F_2}$		Ratio comparison: Experimental / Derived
stock	dilution	stock	dilution	Experimental	Derived	
Sol. 1	PpIX 3 (1%)	Sol. 1	DCM 1 (100%)	97.06	81.52	1.19
			DCM 2 (10%)	16.03	19.01	0.84
			DCM B (8.33%)	15.63	16.06	0.97
			DCM E (4.82%)	10.06	9.70	1.04
			DCM I (2.33%)	5.30	4.84	1.10
			DCM M (1.12%)	2.55	2.37	1.08

Table 5.5 - Comparison of experimental and derived fluorescence ratios for different concentration combinations of PpIX and DCM. Percentage in parentheses represents relative concentration. (Sol.=solution, Prev.=Previous)

Figure 5.17 illustrates the strong correlation between experimental and derived fluorescence ratios, with $R^2 = 0.9900$ when results from all light sources are combined (Figure 5.17a). There is better correlation using the Thorlabs LED light source (Figure 5.17b) compared to the operative light sources (Figure 5.17c and d). This is probably due to the more robust and tested experimental setup using the Thorlabs LED compared to the cumbersome microscope and endoscope experimental setup, where stray light is likely to be a confounding factor despite the use of black-out sheeting. This is reinforced by the better correlation between experimental and predicted values seen for the endoscope compared to the

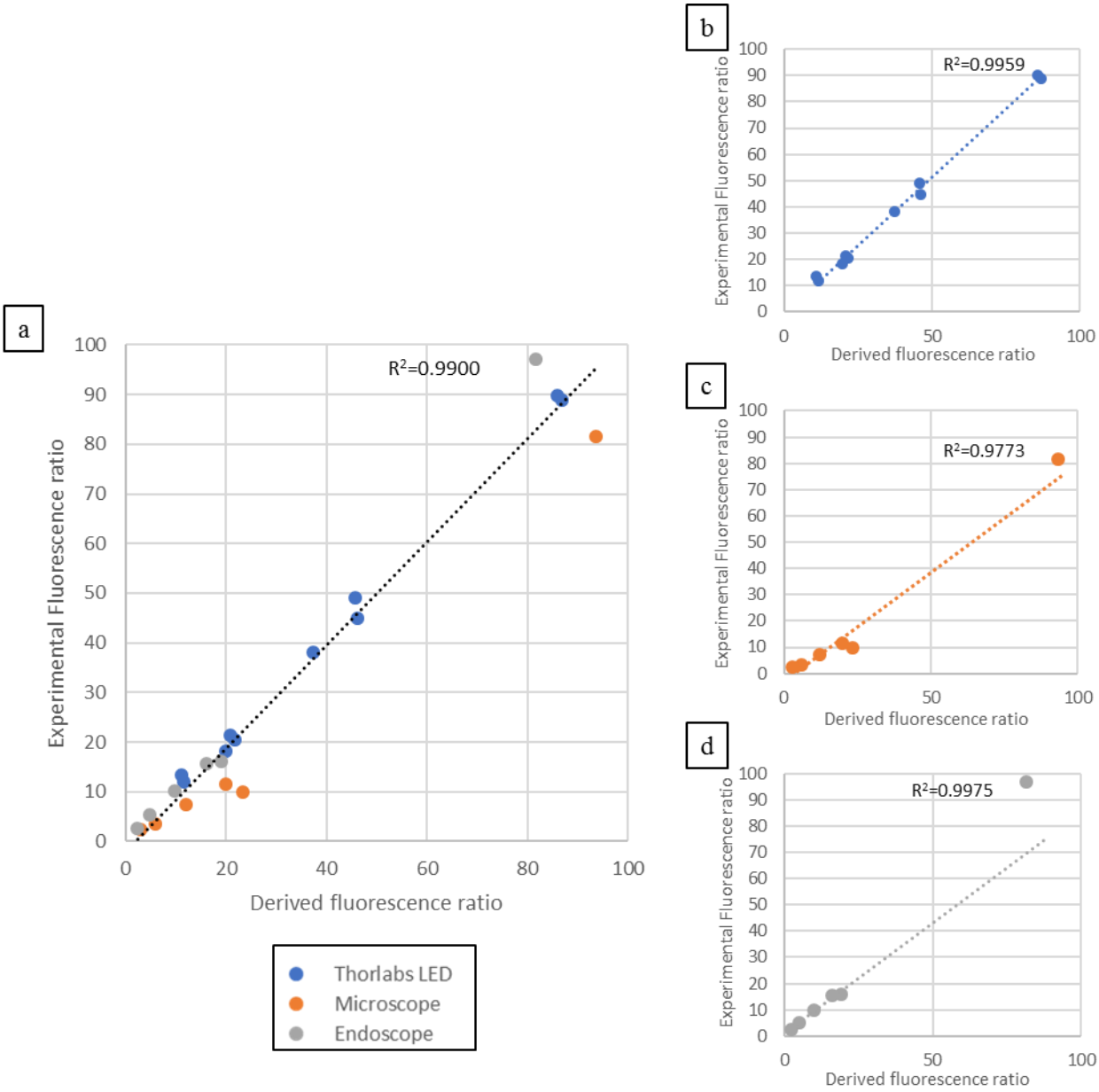


Figure 5.17 - Correlation between experimental and derived fluorescence ratios. Experimental fluorescence ratio values plotted against corresponding derived fluorescence ratio values for a) all light sources, b) Thorlabs LED, c) microscope and d) endoscope. Dotted line represents fitted linear regression model with corresponding R^2 value.

microscope. In the case of the microscope, the working distance is a factor of 10 greater, making the closed experimental setup more difficult to achieve.

5.7.4 Summary and interpretation

The ratio between DCM and PpIX fluorescence has been derived from the Beer-Lambert law and predicted based on measured extinction coefficients for PpIX and DCM. An excellent correlation is demonstrated between the resulting predicted fluorescence ratios and those measured directly through fluorescence measurements on the two samples. DCM can now confidently be used as a surrogate for PpIX in fluorescence experiments. Using equation [11] it is possible to calculate the corresponding PpIX concentration that would give the same fluorescence intensity as any concentration of DCM solution.

5.8 Comparison of microscope and endoscope fluorescence

The use of endoscopic 5ALA-induced fluorescence-guided glioma surgery has been pioneered and developed by our group with only a handful of reports from other centres worldwide[225,226]. During endoscopic fluorescence-guided resection of gliomas it has been observed anecdotally that the surgeon has identified fluorescence when using the endoscope that was not seen through the surgical microscope – the endoscope would therefore appear to offer better visualisation of the PpIX fluorescence. The intensity calculations reported in Section 5.3.2 support this observation, with the endoscope light intensity being greater than the corresponding microscope intensity when the endoscope is positioned at a 20 mm working distance or less. Greater light intensity results in a larger number of excited PpIX molecules, and thus greater fluorescence. This holds true as long as there is no fluorescence saturation at high light intensities, and no evidence of this was found in the saturation experiments reported here (Section 5.6). In the following, we report experiments aimed at quantifying human ability to identify and compare the fluorescence detection threshold for the microscope and endoscope to further substantiate the hypothesis of superior fluorescence visualisation of the endoscope.

Two methods where used to compare the fluorescence detection thresholds, the first using fluorescence spectra and the second using subjective visual detection of fluorescence of serial dilutions of DCM.

5.8.1 Method

Stock solution "DCM solution 1" underwent serial dilutions as shown in Table 5.4. A sample of just the solvent, ethanol (EtOH), was also used as a control. The experimental setup was as outlined previously except that the collection fibre was positioned 110 mm from the surface. This gave a collection area which was larger than the maximum spot size of the light sources and ensured that all fluorescence from the spot was collected. The microscope was positioned at its minimum working distance of 200 mm from the illuminated surface. The endoscope was set up at two different working distances: 30 mm and 10 mm. Both light sources were orientated at 90° to the horizontal. All data were collected with one light source and working distance combination and then the experiment was repeated with a different combination.

For each concentration going down the dilution series a fluorescence spectrum was recorded. The fluorescence intensities were divided by the area of the light source spot to give a value of fluorescence per unit area illuminated. A non-linear curve fitting model was built using normalised DCM and solvent (control) spectra. This method was repeated three times for each concentration and mean curve fit factors were derived and plotted against the relative DCM concentration to give an estimate of the threshold of fluorescence detection of each light source setup.

In addition to the spectrum recording, at each concentration two people viewed the fluorescence directly as a surgeon would in theatre and subjectively reported the presence or absence of fluorescence. For the microscope this was via the operator's eyepieces[xxiii] and for the endoscope this was via a high definition LCD screen. An image of the fluorescence was also acquired using the inbuilt camera of that particular device. These images were cropped to show the same area of solution surrounded by a black background. As an example, a selection of microscope images is shown in Figure 5.18. Six subjects were shown the cropped images for each solution, with the image of the control substance (EtOH) shown between the images of each different concentration of solution in order to eliminate any bias arising from viewing of the previous image. The concentration at which a subject could no longer see fluorescence was recorded

[xxiii] Surgical microscopes often have several eyepieces allowing visualisation by multiple people. The degree of light to each of these is variable but the internal optics are set up to give maximal light, clarity and binocular vision to one main eyepiece known at the operator's eyepiece.

and the subjective threshold of visual fluorescence detection was defined as the lowest concentration at which fluorescence could be seen.

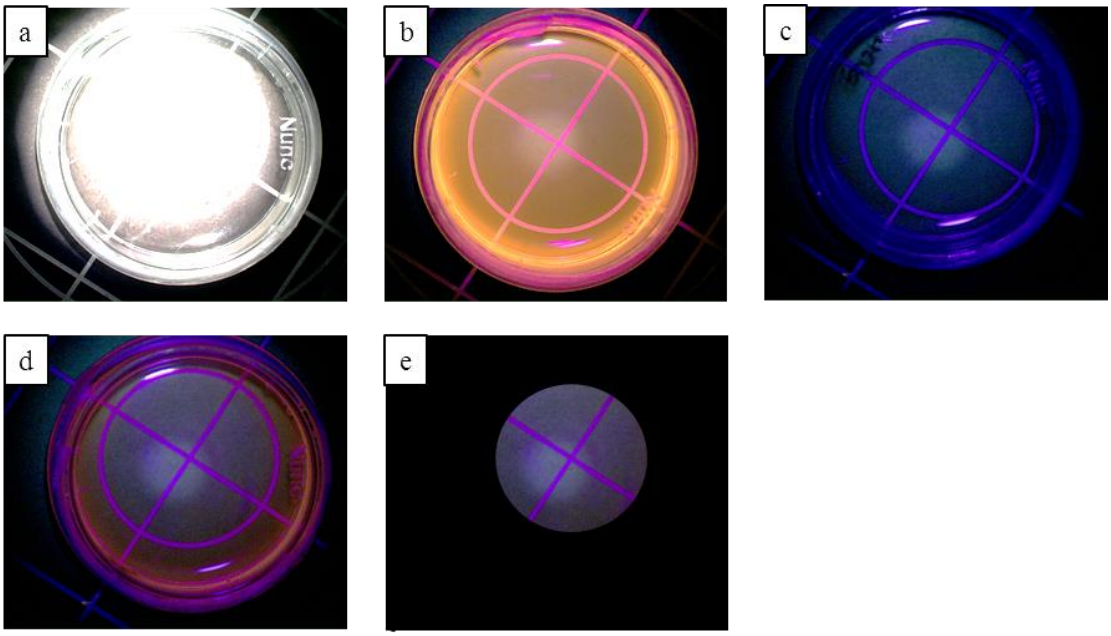


Figure 5.18 – Selection of microscope images. a) DCM 2 illuminated with white light. b) DCM 1 illuminated with blue light. c) Control solution (EtOH) illuminated with blue light. d) DCM 2 illuminated with blue light. e) Example of image cropping using DCM 2 image.

5.8.2 Results

Mirroring the light intensity results in Section 5.3.2, the measured fluorescence intensity is substantially greater when using the endoscope at 10 mm compared with the microscope at 200 mm, but the endoscope results in less fluorescence than the microscope when used at a working distance of 30 mm (Figure 5.19).

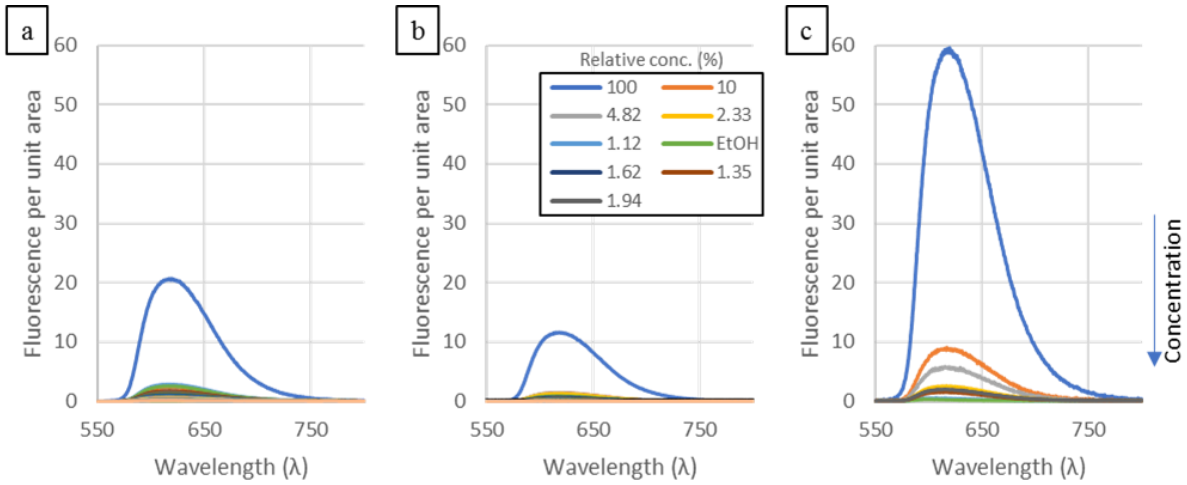


Figure 5.19 - Fluorescence spectra at decreasing DCM concentration. a) Microscope at 200 mm working distance. b) Endoscope at 30 mm working distance. c) Endoscope at 10 mm working distance. Arrow depicts decreasing concentrations with legend showing relative concentrations (relative conc. (%)). Fixed y-axis scale for comparison.

Normalised reference spectra used for the curve fitting model are shown in Figure 5.20 . The relative weights for DCM and solvent determined from the fitting procedure are shown in Figure 5.21. The DCM concentration at which the solvent and DCM contributions to the signal are equal can be regarded as an estimate of the fluorescence detection threshold, and is in the range of 3.20 x 10^{-7} to 3.94 x 10^{-7} mol dm^{-3} (0.65 to 0.8 % relative concentrations) for the microscope, 1.97 x 10^{-7} to 2.71 x 10^{-7} mol dm^{-3} (0.40 to 0.55 % relative concentrations) for the endoscope at 30 mm and 7.40 x 10^{-8} to 1.48 x 10^{-7} mol dm^{-3} (0.15 to 0.3 % relative concentrations) for the endoscope at 10 mm (Figure 5.21b, d and f). The standard deviation error bars in Figure 5.21 demonstrate reassuringly small variance in the data over the three spectra collected at each concentration. Using this model, the endoscope threshold of detection is estimated to be lower than the microscope at 30 mm as well as 10 mm working distance. This does not correspond with the differences in intensity reported above, where one would expect the microscope to perform better than the endoscope at 30 mm given the greater intensity of the former. One explanation might be a case of higher background in the microscope measurements compared to the endoscope, rather than less fluorescence. In the curve fitting model depicted in Figure 5.21 the microscope fluorescence signal is around double that of the endoscope at 30 mm whereas the microscope solvent signal is almost

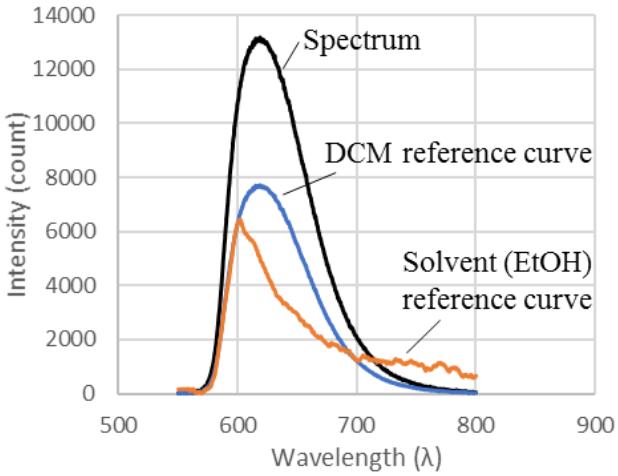


Figure 5.20 - Reference curves used for curve fitting model

four times greater than that with the endoscope at 30 mm (Figure 5.21b and d). The relative signal to noise ratio would therefore be lower in the microscope setup and resulting in a higher limit of detection.

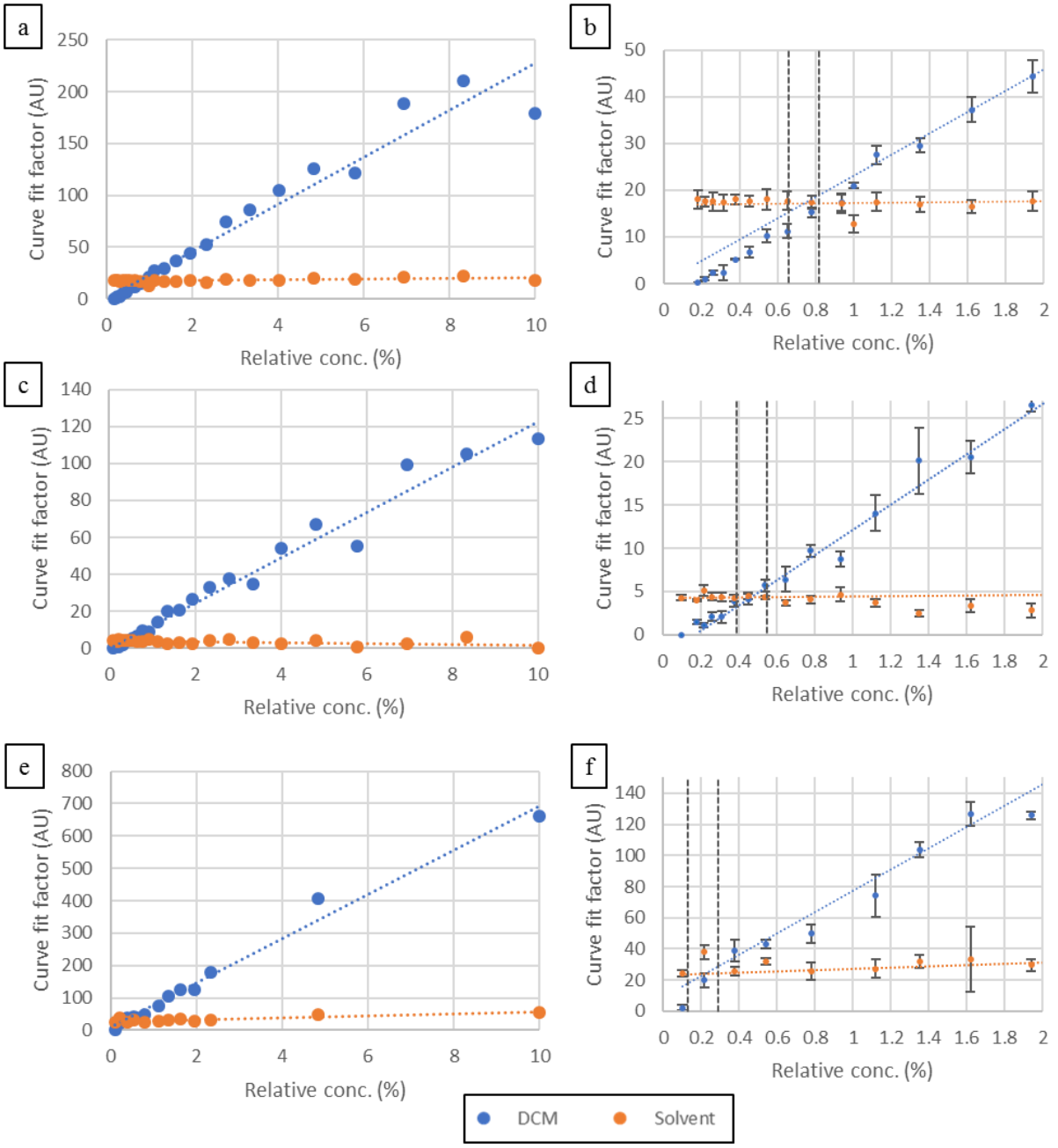


Figure 5.21 - Curve fitting model. DCM and solvent contribution factors are plotted against relative percent concentration (Relative conc. (%)). For illustration purposes two concentration scales are shown 0 to 10 % (left) and 0 to 2% (right). Model factors shown are for the microscope at 200 mm [a) and b)], the endoscope at 30 mm [c) and d)] and the endoscope at 10 mm [e) and f)]. Error bars represent one standard deviation. Dotted lines (blue and orange) represent fitted linear regression model. The concentration range in which the DCM and solvent factors intersect is estimated by the grey hashed lines.

With regard to the subjective visual fluorescence threshold, Figure 5.22 summarises the lowest relative concentration at which subjects reported fluorescence either via direct visualisation via the surgical instrument (eyepiece or high definition LCD screen) or by viewing the fluorescence images taken. Once again, there is a clear advantage of using the endoscope, with fluorescence being detected down to a threshold of almost half the concentration threshold of the microscope. In keeping with the curve fitting

model results above, this advantage is still present, to a certain extent, when the endoscope is at 30 mm working distance. Similar thresholds were found in the direct observation and in the image view which validates the latter as a suitable method of visual fluorescence assessment outside of the operating theatre.

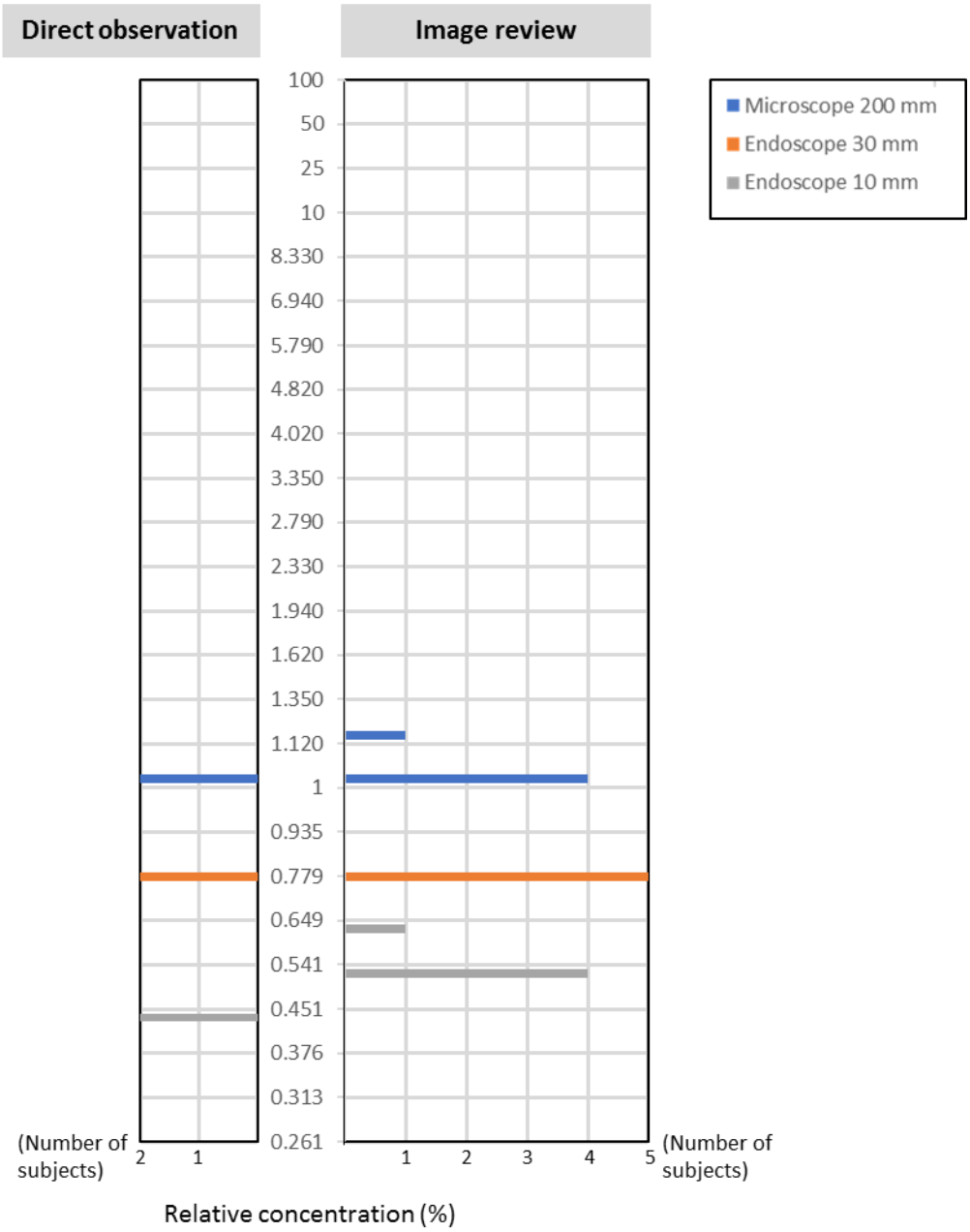


Figure 5.22 - Visual fluorescence threshold. Number of subjects plotted against the last relative concentration of DCM at which fluorescence was seen (visual fluorescence threshold) for both the direct observation via the microscope or endoscope (2 subjects) or on review of the fluorescence images (5 subjects).

5.8.3 Summary and interpretation

These data support the hypothesis that the endoscope provides better visualisation of fluorescence compared to the microscope, with the endoscope having a lower fluorescence detection threshold both

when using a spectrometer to detect the fluorescence and on subjective visual assessment of fluorescence. The DCM concentration detection limits are summarised in Table 5.6. The corresponding concentrations of PpIX that would give the same fluorescence have been calculated using equation [11] (Section 5.7.1) and are also shown in Table 5.6.

Detection method	Operative device	Fluorescence threshold conc. DCM (mol dm-3)	Corresponding conc. PpIX giving same fluorescence (mol dm-3)
Spectroscopic detection	Endoscope at 30 mm	$1.97\text{x}10^{-7}$ - $2.71\text{x}10^{-7}$	**$4.23\text{x}10^{-7}$ - $5.88\text{x}10^{-7}$**
	Endoscope at 10 mm	$7.40\text{x}10^{-8}$ - $1.48\text{x}10^{-7}$	**$1.57\text{x}10^{-7}$ - $3.16\text{x}10^{-7}$**
	Microscope	$3.20\text{x}10^{-7}$ - $3.94\text{x}10^{-7}$	**$8.79\text{x}10^{-7}$ - $1.09\text{x}10^{-6}$**
Visual detection	Endoscope at 30 mm	$3.84\text{x}10^{-7}$	**$8.45\text{x}10^{-7}$**
	Endoscope at 10 mm	$2.22\text{x}10^{-7}$ - $2.67\text{x}10^{-7}$	**$4.79\text{x}10^{-7}$ - $5.79\text{x}10^{-7}$**
	Microscope	$4.93\text{x}10^{-7}$ - $5.52\text{x}10^{-7}$	**$1.39\text{x}10^{-6}$ - $1.57\text{x}10^{-6}$**

Table 5.6 - Fluorescence threshold for DCM and corresponding PpIX conc

Figure 5.23 illustrates the spectroscopic and visual detection thresholds by the endoscope and microscope, in terms of PpIX concentrations (log scale for clarity) and compares them with estimated concentrations of PpIX in high grade glioma and normal brain tissue. These concentration estimates come from a series of papers by Valdes and Wilson *et al.* reporting a method of determining the concentration of PpIX, using a light-transport model which extracts the quantitative fluorescence spectrum from the other tissue optical properties[289,290], in both *ex vivo*[291] and *in vivo*[292] glioma and normal brain samples. The *in vivo* measurements used a fibreoptic probe to interrogate tumour at the time of surgical resection, with normal brain parenchyma, "control tissue", taken from the surgical access tissue. For concentrations in the *in vivo* study, samples were stratified into whether visual fluorescence was seen in the sample or whether fluorescence was only detected by the fluorescence probe, showing no visual fluorescence (labelled "Visual PpIX conc. in glioma" and "Non-visual PpIX conc. in glioma", respectively, in Figure 5.23). Comparing these published concentrations with the PpIX concentration

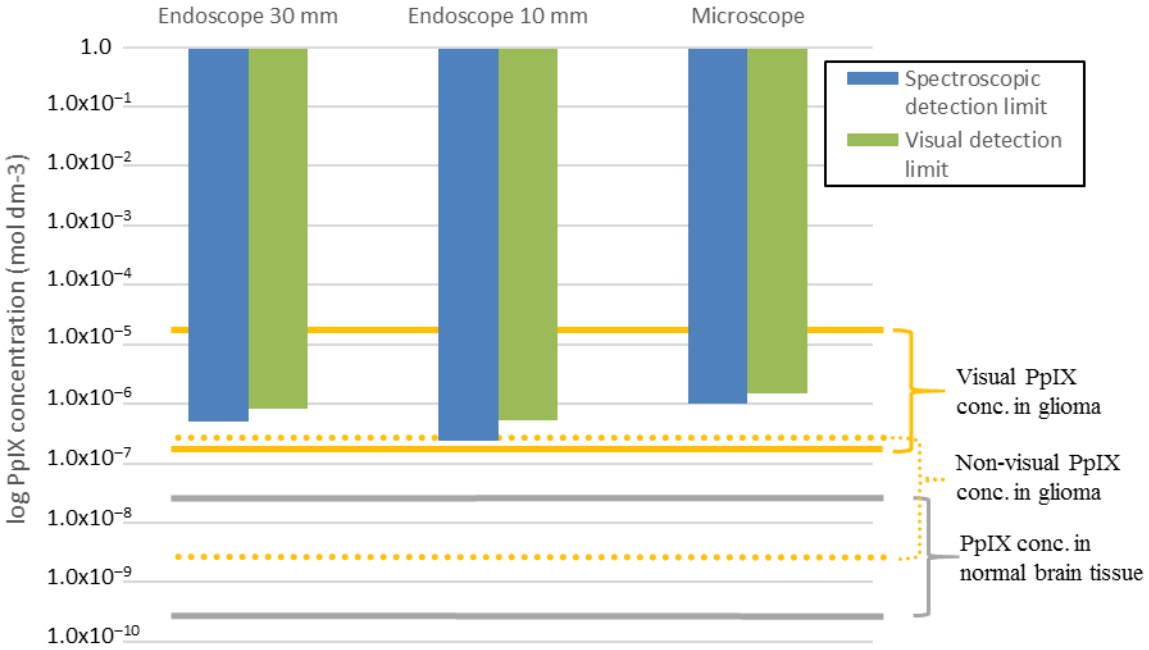


Figure 5.23 - PpIX concentration thresholds and comparison with literature. Spectroscopic (blue) and visual (green) PpIX concentration detection thresholds for endoscope at 30 mm and 10 mm working distance and for microscope. PpIX concentrations (conc.) reported in literature also shown for comparison. log scale is used for clarity.

detection thresholds derived from the work presented in this chapter, it is reassuring that both endoscope and microscope setups have thresholds within the range of visual PpIX concentrations in glioma. Valdes *et al.* used the same model microscope for their study (Zeiss OPMI Pentero) as in this study but seemed to achieve better fluorescence visualisation of fluorescence than in the study presented here. This demonstrates the variability possible in the degree of fluorescence seen which is affected by several factors including, but not limited to, surgical optical devices setup, concentration of 5ALA-in tumour cells, operative theatre background lighting and blood in surgical field. This is a reminder that visual fluorescence is not a quantitative marker and is an argument for using a spectroscopic-based intraoperative tool which would allow for more objective and quantitative measures of the degree of fluorescence and hence concentration of PpIX and amount of tumour. The endoscope, at 10 mm setup in particular, has shown superior levels of detection of PpIX throughout the comparisons presented here but there is still a substantial amount of fluorescence that is not detected both visually and spectroscopically. Spectroscopically, the Valdes *et al.* fluorescence probe setup appears to be significantly more sensitive than the setup reported here at detecting non-visual fluorescence. This is likely to be due to better optimisation for their probe optical setup which would need to be considered if

an intraoperative fluorescence probe was to be considered by our group. Visually, a simple method of increasing endoscope visual fluorescence would be to increase the intensity of the blue light source. The saturation experiments in Section 5.6 confirm that there is unlikely to be a ceiling effect as light intensity is increased. In this context, one should consider possible thermal damage to the tissue as a result of increased light intensity. Little work has been done looking at the effect of thermal damage in neuroendoscopy[293], but the risks are decreased by constant movement of the endoscope and changing the working distance. The blue light intensity could also be increased at the boundaries of the tumour where the PpIX concentration is likely to be reduced.

5.9 Conclusion

Our research group are developing novel methods of enhancing the visualisation of 5ALA-induced fluorescence in order to increase the extent of resection in glioma surgery. It is important to have a robust laboratory-based fluorescence measurement experimental setup that allows for testing of any new technique of devices prior to testing in the challenging, less accessible, operative theatre environment. This chapter has described the development of this laboratory setup. Initially the operative light sources were compared with each other and calibrated with an off-the-shelf light source. The known photodegradation of PpIX was also investigated to establish the practicalities of using PpIX solutions in the laboratory setting. The laser dye DCM was proposed as a stable alternative solution possessing very similar absorption and fluorescence properties as PpIX. DCM fluorescence was then correlated with PpIX fluorescence, using both a derived method from first principles and transmission spectra, and an experimental method of direct fluorescence measurements. The correlation between the derived and experimental parameters was reassuringly high.

Once the setup was established it was used to compare the operative microscope with the endoscope at different working distances showing both visually and spectroscopically that the endoscope outperforms the microscope in terms of its ability to show fluorescence. From a surgical point of view, it is important to remember that the closer the endoscope light source is to the area of interest, the more fluorescence will be detected. The endoscope tip needs to be within 30 mm, or ideally 10 mm, distance from the surface for the endoscope to have a benefit over using the microscope.

This established experimental setup will now be used to test further methods of enhancing fluorescence detection. The suggestion in the previous section of increasing the endoscope blue light source intensity could be investigated, and our group are in discussion with Karl Storz regarding testing new blue light sources. Development of an intraoperative fluorescence probe will also be carried out on this laboratory setup.

Chapter 6. Identification of tumour versus normal brain using Raman spectroscopy

6.1 Introduction

The development and applications of Raman spectroscopy and fluorescence spectrometry reported in this thesis so far have been carried out in parallel. The most relevant intraoperative application of both techniques in neuro-oncology surgery is the identification of the tumour transition zone and differentiating between tumour and normal brain. The future direction of this project will take a multimodal approach, trialling many novel methods of enhanced tumour detection, including Raman spectroscopy, fluorescence spectrometry and other fluorescence-enhanced visualisation techniques. In the case of HGG these will also need to be compared to the gold-standard of 5ALA-induced fluorescence guided surgery. Towards this goal, a preliminary study comparing Raman spectroscopy with 5ALA-induced visual fluorescence is reported in this Chapter.

Our neurosurgery group is particularly well placed to test these modalities intraoperatively as we have established and advocate the use of functional-guided resection of HGG (see Section 1.4.4)[104]. This means the resection routinely goes beyond the visual fluorescence boundary, if safe to do so, and thus there is a region of non-fluorescing tumour resection where these novel tumour-detecting modalities can be evaluated. When functional guided resection allows for tumour resection beyond the visual fluorescence boundary in all planes, and this is confirmed on postoperative MRI, the resection is said to be a "supramaximal", rather than a gross total, resection.

In this preliminary study, fresh tissue biopsies from tumour resection cavities, after completion of functional guided tumour resection, underwent Raman analysis. The main aim was to evaluate the ability of Raman spectroscopy to differentiate infiltrating tumour from normal brain at the extreme margins of the tumour and to compare its predictive performance with the current gold standard of 5ALA-induced fluorescence.

6.2 Method

6.2.1 Study participants and intraoperative biopsy

The study participants were a subset of patients from the genetic classification study (Chapter 4), who underwent functional guided tumour resection where the functional boundaries of resection went beyond the visual fluorescence and the resection was defined as a supramaximal resection. At the end of the surgical resection, the last pieces of tissue removed from the resection cavity were sent as biopsies. The location of each resection cavity biopsy was recorded using the intraoperative neuronavigation system (Brainlab, Germany) (Figure 6.1b) and labelled according to the anatomical location relative to the resection cavity (e.g. anterior). Prior to resection, each sample location was inspected for 5ALA-induced fluorescence using either the operative microscope or endoscope blue light sources (described in Section 5.3.1) and the presence of visual fluorescence recorded. A micro-biopsy forceps were used to take the biopsy, which was immediately taken to neuropathology. After the operation, the routine postoperative MRI scan was reviewed to confirm that supramaximal resection had been achieved[xxiv] (Figure 6.1c).

[xxiv] A volumetric analysis of pre- and postoperative MRI scans has been developed and validated by our group to differentiate supramaximal resection from gross-total resection[107].

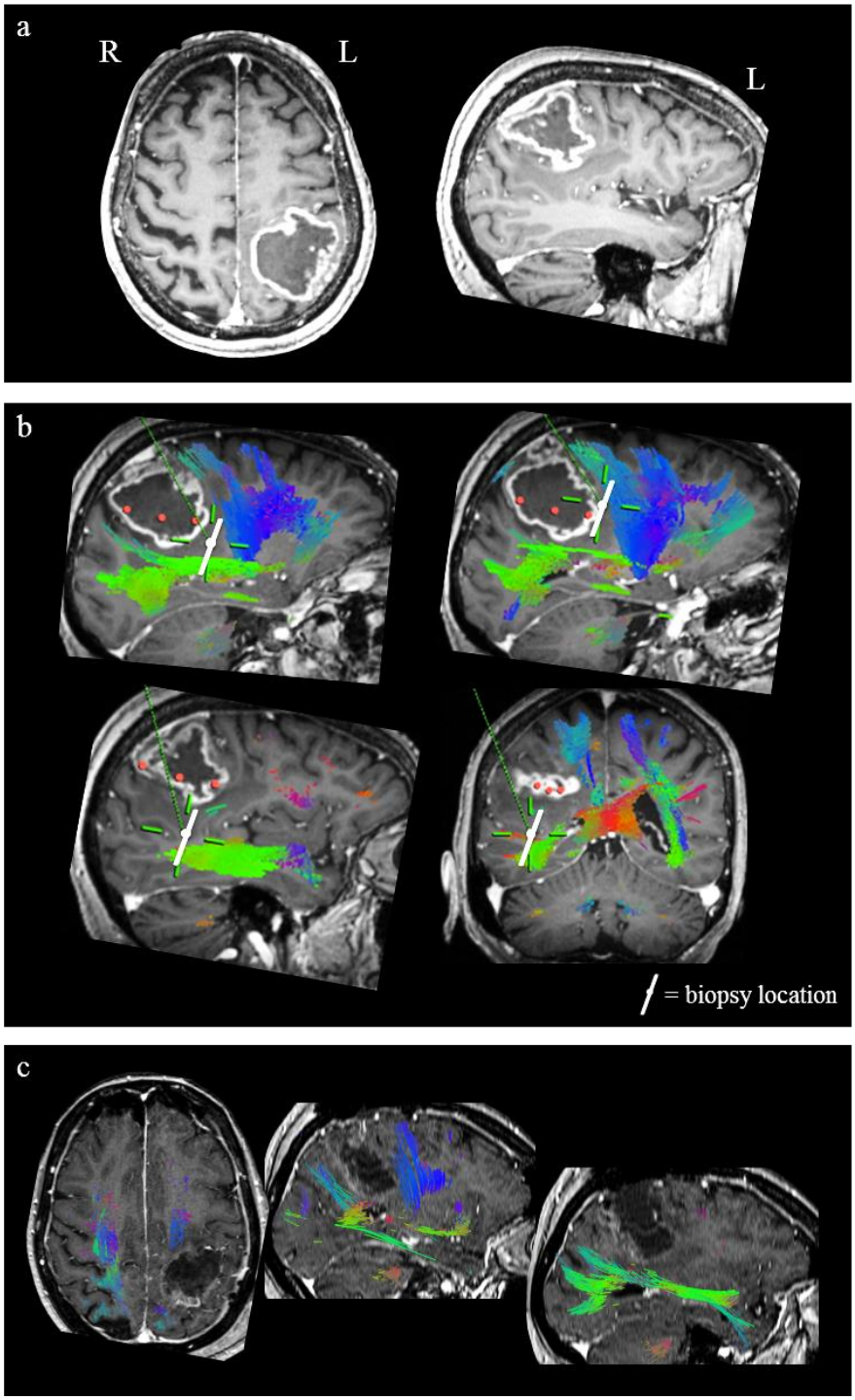


Figure 6.1 – MRI imaging for supramaximal resection. a) Preoperative MRI demonstrating enhancing tumour consistent with glioblastoma. b) Intraoperative neuronavigation showing location of supramaximal resection margin biopsies beyond contrast-enhancing tumour. c) Postoperative MRI demonstrating supramaximal resection up to white matter tracts. All images are T1-weighted sequences with gadolinium contrast. Coloured overlay represents estimated location of white fibre tracts using diffusion tensor imaging (DTI).

6.2.2 Administration of 5ALA

A 20 mg/kg oral dose of 5ALA dissolved in water was given to all patient with suspected glioblastoma on preoperative imaging unless contraindicated. The 5ALA was given 2 hours before the start of surgery in order for peak fluorescence, at around 6 hours[124,146], to coincide with the last stages of tumour resection at the supramaximal boundary.

6.2.3 Normal fresh brain biopsies

In order to build a classification model to differentiate between tumour and normal brain it was necessary to record Raman spectra from several normal fresh brain biopsies. These were acquired from patients undergoing mesial temporal lobectomy for epilepsy. A small amount of temporal brain tissue is routinely removed in order to access the mesial temporal lobe. This access tissue would normally be discarded at the time of surgery, but can be collected as a source of normal fresh brain tissue and used for research. In selected cases large biopsies (approximately 10 mm^3 in volume), incorporating both cortical grey and white matter, were taken. Part of the biopsy underwent standard H&E staining and neuropathology evaluation to confirm the presence of normal brain tissue, the remaining sample underwent immediate Raman analysis using the fresh tissue analysis method outlined in Chapter 4.

6.2.4 Neuropathological evaluation

Each biopsy sample, approximately 3 mm^3 in volume, was cut in half by the neuropathologist. One half was taken straight for Raman analysis, the other half was processed as part of the routine diagnostic pathway undergoing formalin fixation and paraffin embedding. The latter sample underwent standard H&E staining. Depending on the pathological diagnosis of the main tumour samples from the case, further IHC tests were carried out. If the main tumour samples resulted in a diagnosis of an IDH-mutated tumour (astrocytomas, IDH-mutant; glioblastoma, IDH-mutant; oligodendroglioma) then the resection margin biopsies underwent IDH-R132H IHC to identify tumour cells within the sample. If the main tumour samples lead to a diagnosis of IDH-wildtype tumour (astrocytoma, IDH-wildtype; glioblastoma, IDH-wildtype), then the margin biopsies underwent Ki-67/MIB1 IHC. MIB is an antibody for Ki-67 protein which is expressed in the active phase of the cell cycle, thus Ki-67/MIB1 is a marker of cell proliferation[294]. Although not specific for tumour cells, the amount of cell proliferation in a section gives

an indication of the presence of tumour infiltration. With the combination of the standard H&E and the appropriate IHC stain, the neuropathologist categorised the margin biopsies into 3 groups:

1. **Tumour** – sample is characteristic of tumour throughout.
2. **Infiltrating tumour** – evidence of tumour cell infiltration into normal white or grey matter (the degree of infiltration can vary but was not sub-divided in this study).
3. **No tumour cells** – histological and IHC features of normal brain with no evidence of tumour.

6.2.5 Raman spectroscopy

Raman spectroscopy was performed on the fresh resection margin biopsy samples immediately after the biopsy was cut by the neuropathologist. The same method as outlined in Chapter 4, for the fresh tissue analysis in the genetic classification study, was used. The margin biopsy samples were generally smaller than the samples in the genetic classification study, and therefore the aim was to obtain at least 50-100 spectra per fresh tissue sample.

6.2.6 Classification model and statistical analysis

A two-group classification model – tumour versus normal brain – was built using PCA-LDA as outlined in Chapter 4. For the "tumour" group, the input spectra came from the fresh tissue spectra from almost all tumour samples included in the genetic classification study, regardless of genetic or pathological subtype. To ensure that the model was independent from the test data, tumour spectra from cases where resection margin biopsies were taken were not included in the training set. The "normal brain" group input spectra were taken from the fresh tissue analysis of normal brain samples acquired from epilepsy surgery cases as described above.

As described in Chapter 4, ANOVA was used to select the PCs which show a statistical difference between the two classification groups. Model performance was assessed using leave-one-patient-out cross-validation to generate model sensitivity and specificity. All spectra from resection margin and normal brain biopsies underwent pre-processing as outlined in Chapter 4. The mean spectra and the second derivative of the mean spectra were calculated. The Wilcoxon rank sum test was used to test for statistical significance between the intensities of the spectra in each group.

6.2.7 Classification of resection margin spectra

Spectra from the resection margin samples were individually classified into one of the two groups in the built model (individual spectra classification). The percentage of spectra classified into each group for a given biopsy was also calculated, and the group with the greatest number of spectra classified into it was taken to be the resultant overall classification of that biopsy (overall biopsy classification). Values of sensitivity, specificity and accuracy were calculated in each case, along with 95% confidence interval calculations (Clopper-Pearson confidence intervals[295-297]) and compared with similar parameters for 5ALA-induced fluorescence where applicable. McNemar's test for related categorical data[298,299] was used to measure the statistical differences between Raman and 5ALA for differentiating tumour versus normal brain.

6.3 Results

Raman analysis of 23 supramaximal resection margin biopsies was undertaken in a total of eight patients. The cases and corresponding biopsies are summarised in Table 6.1. There were nine biopsies from three glioblastoma, IDH-wildtype tumours, eight biopsies from three glioblastoma, IDH-mutated tumours and six from two oligodendrogliomas (one WHO grade II and one WHO grade III). The neuropathology assessment categorised two biopsies as tumour, 17 as infiltration zone and four as no tumour cells.

Eleven fresh normal brain biopsies were acquired from 11 patients undergoing temporal lobectomy for epilepsy. Neuropathology review confirmed normal grey and white matter in every case. After pre-processing, 1825 spectra of fresh normal brain were used in the model, with a mean of 166 spectra per sample. For the tumour spectra, the previously descibed 9,799 spectra from 62 fresh tissue samples used in the genetic classification study were used (Chapter 4). The model performed extremely well when assessed using leave-one-patient-out cross validation, as shown by the high sensitivity and specificity illustrated in Figure 6.2 ,which also shows the LD score plot for the model. The mean spectra and the negative of the second derivative of the mean spectra for tumour and normal brain are illustrated in Figure 6.3. The difference spectrum (tumour minus normal brain) is also shown in this figure. The prominent peaks are similar to those observed in Chapter 4, and the detailed peak allocation presented in Chapter 4 can also be applied here. Lipid peaks are more prominent in tumour compared to normal brain

spectra, with dominant peaks at 825 cm^{-1}, 853 cm^{-1}, 1087 cm^{-1}, 1124 cm^{-1} and 1305 cm^{-1}. This is also true for peaks representing DNA (517 cm^{-1}, 885 cm^{-1}, 1206 cm^{-1} and 1342 cm^{-1}). Conversely, normal brain spectra have stronger amino acid peaks (540 cm^{-1}, 615 cm^{-1}, 635 cm^{-1} and 649 cm^{-1}) and protein peaks (1522 cm^{-1} and 1553 cm^{-1}, both representing Amide III protein) compared with tumour spectra. 5ALA was used in all glioblastoma cases, and fluorescence was observed at the site of biopsy in one out of 17 biopsies (Table 6.1). For the cases in which 5ALA was used, absence of fluorescence was highly specific (specificity=1.00, rate of true negatives), as in all biopsies where there was no tumour there was also no fluorescence, but poorly sensitive (sensitivity=0.07) as most biopsies containing tumour did not fluoresce.

Pathology (WHO grade)	Resection margin biopsy location	5-ALA fluorescence seen at location of quadrant?	Neuropathology assessment			Classification of spectra		
			IDH	MIB-1	Categorisation	Number of spectra*	Proportion classified as TUMOUR	Proportion classified as NORMAL
GB-WT	Anterior	None		0	No tumour cells	65	0.37	0.63
	Posterior	None		5%	Infiltrating tumour	194	0.99	0.01
GB-WT	Posterior	None		10%	Infiltrating tumour	217	0.99	0.01
	Inferior	Faint fluorescence		5%	Infiltrating tumour	163	0.96	0.04
	Posterior Inferior	None		10%	Infiltrating tumour	160	0.92	0.08
	Medial	None		10%	Infiltrating tumour	180	0.94	0.06
GB-WT	Anterior	None		5%	Infiltrating tumour	99	1	0
	Medial	None		5%	Infiltrating tumour	160	0.95	0.05
	Posterior	None		5%	Infiltrating tumour	122	0.99	0.01
GB-MUT	Inferior	None	+		Tumour	144	1	0
	Posterior	None	+		Infiltrating tumour	183	0.97	0.03
	Superior	None	+		Infiltrating tumour	105	1	0
GB-MUT	Anterior	None	+		Tumour	107	0.97	0.03
	Posterior	None	-		No tumour cells	206	0.23	0.77
GB-MUT	Inferior	None	+		Infiltrating tumour	86	0.94	0.06
	Posterior	None	-		No tumour cells	183	0.31	0.69
	Medial	None	+		Infiltrating tumour	108	1	0
Oligo (III)	Posterior	Not used	+		Infiltrating tumour	72	0.97	0.03
	Inferior lateral	Not used	+		Infiltrating tumour	42	0.83	0.17
	Superior	Not used	+		Infiltrating tumour	186	0.97	0.03
Oligo (II)	Frontal	Not used	+		Infiltrating tumour	78	0.97	0.03
	Temporal	Not used	-		No tumour cells	170	0.41	0.59
	Posterior	Not used	+		Infiltrating tumour	88	1	0

*Table 6.1 – Supramaximal resection cases and resection margin biopsies. Shaded cell indicates the overall classification of each biopsy. IDH + and – refers to sample staining positive for IDH R132H mutation (+) or not (-). MIB-1 score given as a percentage of cells per field of view. *Number of spectra after pre-processing.*

Classification group	Sensitivity	Specificity	Accuracy
Tumour	0.96	0.99	0.99
Normal brain	0.99	0.96	0.99

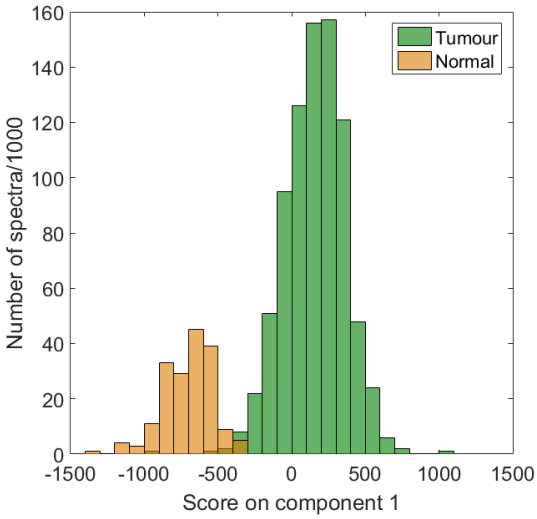


Figure 6.2 – PC-LDA model of tumour versus normal brain. Performance of model using a leave-one-patient-out cross validation (left). LDA score histogram (right).

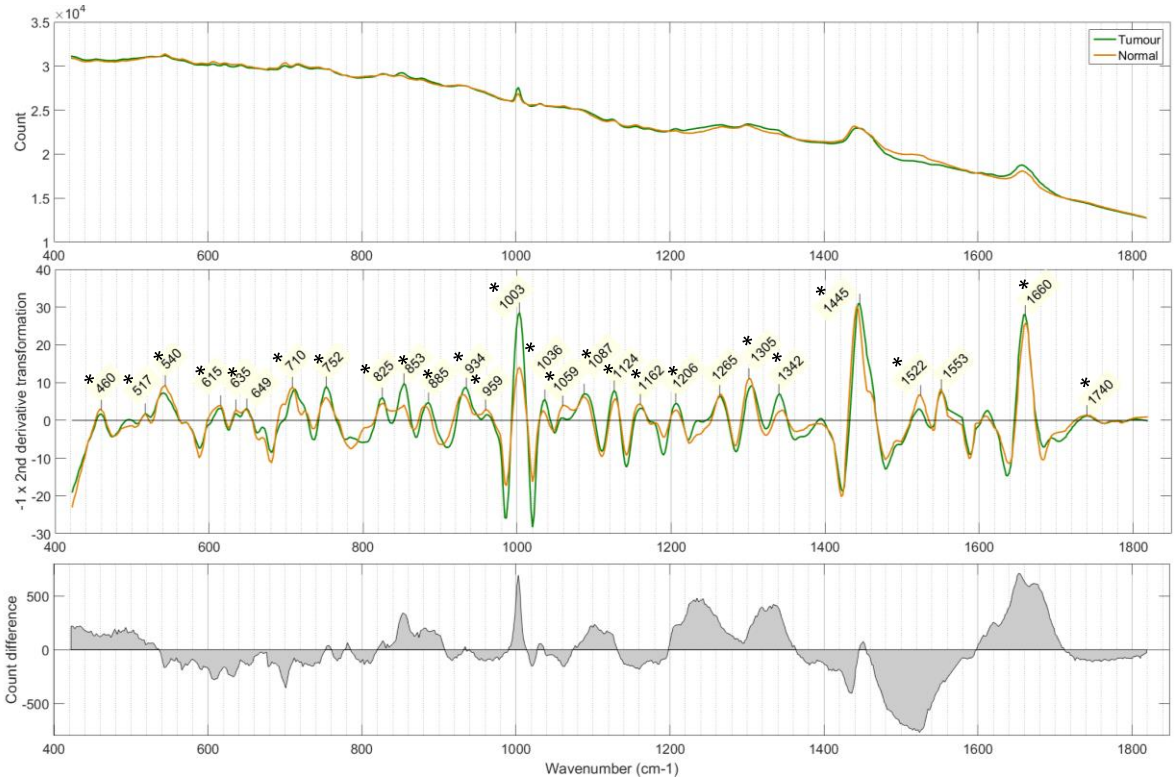


*Figure 6.3 - Tumour and Normal brain spectra. Mean spectra (top), difference spectra (bottom), and negative of second derivative transformation of mean spectra for tumour and normal brain after EMSC (centre). The wavenumber of prominent Raman peaks are labelled and the statistical significance ($p < 0.01$) of the difference between tumour and normal brain peak intensities is indicated by *.*

Table 6.2 compares the performance of 5ALA-induced fluorescence for predicting tumour versus normal brain with the performance of Raman classification models. Raman performance depends on which method is used to classify the biopsy samples into tumour/normal brain groups. Using overall biopsy

classification – classifying the biopsy based on the group with the greatest percentage of spectra classified to it – the performance is extremely high, with correct classification of biopsies in every case, although there is a wide confidence interval range due to the low *n*. This performance is significantly better than 5ALA-induced fluorescence ($\chi^2 = 11.077$, $p = 0.0009$). Using individual spectra classification, there is still far superior, and statistically significant, performance compared to 5ALA-induced fluorescence ($\chi^2 = 1920.12$, $p = 0.0001$). The performance is also high when oligodendroglioma cases

Modality (classification method (*n*))	**Sensitivity**	**Specificity**	**Accuracy**
5-ALA induced fluorescence (17)	0.07 (0.02 to 0.34)	1.00 (0.29 to 1.00)	0.24 (0.07 to 0.50)
Raman (overall biopsy classification, either including (23) or excluding oligo (17))	1.00 (0.77 to 1.00)	1.00 (0.29 to 1.00)	1.00 (0.85 to 1.00)
Raman (individual spectra classification (2482))	0.97 (0.96 to 0.98)	0.72 (0.67 to 0.76)	0.93 (0.91 to 0.94)
Raman (individual spectra classification including oligo (3118))	0.97 (0.96 to 0.98)	0.68 (0.64 to 0.72)	0.91 (0.90 to 0.92)

Table 6.2 - Performance of 5-ALA and Raman to detect tumour versus normal brain. Range in parentheses gives 95% confidence intervals. (oligo=oligodendroglioma)

are included.

6.4 Discussion

This preliminary study shows that Raman spectroscopy can accurately differentiate between infiltrating tumour and normal brain at the margin of a functional guided, supramaximal resection cavity. Raman spectroscopy is also significantly superior to 5ALA-induced fluorescence in predicting tumour versus normal brain at this margin.

In the bulk of the tumour and up to the contrast-enhancing margin seen on image guidance there is little doubt that 5ALA-induced fluorescence is a powerful surgical adjunct. The correlation between contrast-enhancing tumour on preoperative MRI and 5ALA-induced fluorescence is well recognised[131], although fluorescence has been shown to be more sensitive than the contrast enhancement in identifying tumour

in some cases[140]. It is this correlation which accounts for increased extent of tumour resection, and progression free survival, when using 5ALA-induced fluorescence guidance surgery[73]. Fluorescence-guidance surgery is conducive to the neurosurgical workflow, only requiring the surgeon to press a button to switch from white to blue light while viewing the tissue through the surgical microscope. There is no need to pause surgical resection or to change instruments as would be the case with a probe-based system. This study focuses on biopsies taken at the boundaries of supramaximal resections, beyond the contrast-enhancement. Here the predictive value of 5ALA is reduced, and there are reports in the literature of between 25 to 65% false negative rate in biopsies taken beyond the visual fluorescent margin[131,134,135,217,291]. This is even lower in the current study, with only one out of 14 biopsies with infiltrated tumour showing fluorescence. At the limit of the surgical resection, where there is a low level of tumour infiltration, the decision between further resection balanced against increased risk of neurological disability is most critical and can have the most serious implications for the patient. It is at this limit where a probe-based Raman device would have significant application, showing significantly better predictive value than 5ALA fluorescence. It is perhaps disheartening that in this study most supramaximal resection margin biopsies were found to have infiltrating tumour, but it should not be surprising – these are, after all, named "*diffuse* gliomas" by the WHO[300]. Despite this, previous research by our group has shown that supramaximal resection in glioblastoma results in a statistically significant increase in progression free survival compared to gross total resection[104]. The need for tumour versus normal brain differentiation at the margins of these resections is therefore needed.

The principal reason that Raman spectroscopy outperforms 5ALA fluorescence imaging is the higher detection rate, i.e. higher sensitivity, for tumour infiltration, with over 95% of spectra of infiltrating tumour correctly classified in most biopsy samples. The correct classification of normal brain is less accurate, with between 59 and 77% of spectra from samples with no tumour cells correctly classified. Contrary to this, the classification model built to classify the margin biopsies showed high sensitivity and specificity on leave-one-patient-out internal cross validation. Here the normal brain spectra came from samples from patients with epilepsy and no tumour. The biochemical composition, and thus Raman spectra, of this normal brain tissue may differ from the resection margin biopsies, which may be affected by the presence of a nearby tumour causing localised raised intra-cranial pressure or peritumoural

oedema. The use of normal brain tissue from non-tumour patients is still valid, and indeed provides the only guarantee that there are no occult tumour cells within the samples. Reassuringly, tumour and normal brain mean Raman spectra derived from the model have similar peak characteristics and differences to those reported in published data on tumour versus normal brain Raman analysis[182-184,190,301]. Moreover, in addition to the data reported here, tumour versus normal brain models were built using the cryosection and FFPE samples from the genetic classification study, and similar spectral differences to those in the fresh tissue model were observed (data not included). As part of the development and validation of an intraoperative Raman probe device, the specificity and false positive rate would need to be fully evaluated, as a high false positive rate may result in the unnecessary removal of normal brain. In Jermyn *et al.*, the only published report of using an intraoperative Raman probe for glioma detection (see Section 1.6.5), the specificity is 91%[190]. This is an encouragingly high figure, but practically means that just under one in 10 areas that are sampled will give a false positive result – indicating tumour in an area of potentially normal brain. Jermyn *et al.* explored in more detail the false negative rate and estimated the threshold of detection for their intraoperative device to be approximately 15% cancer cell burden or 17 human cancer cells/0.0625 mm^2. The accuracy of the estimation must be questioned given the methodology used in the study. Raman analysis was carried out *in vivo* (0.5 mm diameter sampling area with a sampling depth of ~ 1mm[xxv189]) and then a biopsy (reported size of ~0.5 mm by ~0.5 mm by ~ 3 mm deep) of the area was taken guided by the indentation left in the brain by the Raman probe. Using this method, the accuracy of the correlation between the area interrogated by the probe and the area reviewed by the neuropathologist is unlikely to be high enough to make accurate conclusions based on only a few cancer cells per histological sample. To address this issue, future intraoperative studies should include an additional step in the protocol where, after intraoperative Raman analysis and biopsy, the fresh sample immediately undergoes ex vivo Raman analysis with the probe device. These data can be correlated with the intraoperative spectra to confirm correlation between intraoperative measurements and sample biopsies. The data from the study in this chapter show that Raman spectroscopy is more sensitive at detecting tumour, and has a lower false negative rate, than 5ALA fluorescence, but that the

[xxv] This was determined using Monte-Carlo based light transport simulations which estimated that ~95% of the probe excitation laser photons were absorbed by 1 mm depth.

false positive rate using Raman analysis might be problematic. Therefore, although accurately establishing the threshold for tumour detection by a Raman probe is an essential part of any future work, the focus should also be on accurately quantifying the false positive rate. Future intraoperative Raman probe projects require the clinical spectroscopy and neurosurgery communities to agree on what level of accuracy is both technically achievable and clinically useful. Practical strategies need to be discussed between the two groups. Perhaps, for example, sampling a brain area multiple times would decrease the risk of a false positive result, but the feasibility of this within the neurosurgical work flow needs to be considered.

The main limitations of this study are the small number of cases and the ex vivo Raman analysis. The objective for the study was to provide proof of concept in an ex vivo Raman study before an intraoperative device would be developed. This objective has been met, and the study has also shown the superiority of Raman over 5ALA. *In vivo* studies will have significant challenges, including reduced performance due to blood or fluid in the measurement field, signal loss from the probe fibre, and noise from surrounding light sources. Some of these issues have been addressed by Jermyn *et al.*[189,190], and some methodological considerations have been mentioned above. Future intraoperative Raman probe evaluation will require significant collaboration between neurosurgeons and clinical spectroscopists in order to produce a usable intraoperative device that is highly accurate and user-friendly, and which could ultimately lead to maximal glioma resection and increase patients' survival.

Chapter 7. Conclusions and future work

This thesis has reported on the application of Raman spectroscopy in glioma surgery and neuropathological diagnosis and the development of a fluorescence spectrometry setup to allow evaluation of future fluorescence-enhanced visualisation and detection techniques.

7.1 Optimisation of Raman spectroscopy experimental protocol

7.1.1 Conclusions

The development of sample preparation for Raman analysis was reported in Chapter 2. In general, is was possible to record good quality spectra of normal and pathological tissue from cryosections and FFPE sections. The main advantage of cryosections is the unprocessed nature of the tissue and the lack of paraffin artefact, which was seen in some FFPE sections particularly when using CaF_2 slides. FFPE sections, however, have a significantly longer life span with no significant change in spectral characteristics even 5 weeks after sample preparation, as reported in Chapter 3. The evaluation of substrates favoured SS over CaF_2 or SiO_2 with better SNR and less substrate background. Sample adherence to the slides was initially a problem, particularly for the CaF_2, which was overcome by the development and use of a slide coating protocol. Practically, SS is cheaper and stronger than CaF_2 lending itself to clinical and intraoperative applications. Laser excitation wavelength was also investigated, with minimal difference in the spectral quality between 830 nm to 785 nm wavelength, reflecting the use of both wavelengths for Raman analysis of biological material in the literature[164]. 785 nm was chosen for the remainder of the project due to the slightly better classification model performance and the availability of a benchtop spectrometer already configured with a 785 nm laser.

In Chapter 3, a number of variables were evaluated to investigate the possible causes for a discrepancy seen between preliminary genetic classification data collected using different laser exposure times (4 s versus 6 s). It was shown that there was no degradation in spectral sample quality over a 5 week period. There was also no significant effect of laser exposure on surrounding tissue. This led to the notion that increasing laser exposure had an effect on the Raman spectral characteristics of the illuminated tissue. Allan variance was used to establish the optimum balance between preventing laser-induced changes in the spectra and acquiring meaningful spectra with sufficient SNR. The phenomenon of background

suppression due to repeat laser exposure was observed throughout the protocol development stage and was characterised as part of the experiments in Chapter 3. The origin of this background was discussed and the findings in this thesis broadly support the hypothesis that the background is due to light scattering rather than tissue fluorescence, with repeated laser exposures causing an annealing effect rather than a quenching of fluorescence.

7.1.2 Future work

The vibrational biospectroscopy research community have made efforts to standardise sample preparation and spectrometer setup for use in clinical applications[209,228]. There are a whole range of methodological and instrumentational variables that could be further investigated and optimised to improve the Raman analysis of biological samples. It is essential that experimental protocols are developed with the end clinical application in mind. It is also important for the biospectroscopy community to promote the accurate reporting of methodology and experimental design to allow replication by others. The reporting of negative results using failed experimental methods should be encouraged to prevent others from wasting research time and money repeating these futile techniques

The strong background signal, and its suppression with repeated exposures, merits further investigation. This project has taken a practical approach, measuring the effect of the background on SNR and classification model performance, and using Allan variance to attempt to gauge the exposure time which is less likely to result in a change in spectral characteristics. Further work needs to focus on the physical origin of this background in order to find practical solutions for background suppression. As mentioned in Section 3.5, using mass spectrometry, to characterise the proteomic and metabolic changes, and electron microscopy, to visualise structural changes on the tissue surface, with varying degrees of laser exposure may provide more insight into the biochemical and physical processes occurring in background suppression.

7.2 Rapid molecular genetic classification of gliomas using Raman spectroscopy

7.2.1 Conclusions

The main clinical application of Raman spectroscopy investigated in this thesis was the rapid molecular genetic classification of gliomas with particular focus on intraoperative, *ex vivo*, analysis of fresh tissue.

A fresh tissue workflow was developed which could easily be followed by the scrub nurses or neurosurgeon, thus allowing for rapid Raman analysis of the tissue without the need for laboratory and neuropathology support. Using a supervised machine learning model with independent cross validation it was possible to use Raman spectroscopy to classify the common glioma genetic subtypes with high accuracy. This was also possible, although with lower accuracy, using cryosection, FFPE sections and glioma cell lines, suggesting possible applications within the neuropathology laboratory, particularly in identifying rare IDH mutations which return a false negative on routine IHC and require genetic sequencing. The tentative biomolecular peak allocations show differences between the genetic subtypes which are consistent with what is understood about the biochemical processes within these subtypes. This, and the consistency between the dominant peaks across all sample preparation types, further validates the technique.

7.2.2 Future work

The limitations of the genetic classification study are reported in Chapter 4. Further work must involve increasing the dataset to include larger numbers of each tumour subtype in order to account for intra-subtype heterogeneity. Data must be collected from multiple centres using different spectrometers to evaluate the inter-user and inter-machine variability. Although the cross validation used in the study was independent, using leave-one-patient-out validation, it would be important to carry out independent valuation of the established model using a new cohort of patients acting as a validation dataset. If classification performance remains high after this, then the technique is ready for a trial of clinical application. After clinical application, ongoing model development and quality control should be carried out by the central collection of all spectra, as they are generated, along with the corresponding final neuropathological diagnosis.

This thesis has focused on the classification of glioma genetic subtypes, which has significant implications for intraoperative decision making, as outlined in Sections 1.4.3 and 4.4. As the treatments for these tumour subtypes are refined, and the subtypes themselves are redefined, the Raman diagnostic models will need to be adapted. It will be important for the biospectroscopy community to continue to identify the important intraoperative and clinical questions, and how vibrational spectroscopy techniques can help to answer them.

7.3 Development of laboratory-based fluorescence spectrometry setup

7.3.1 Conclusions

Chapter 5 reported on the initial stages of development of fluorescence spectrometry applications in glioma surgery in our research group. The development of a laboratory-based fluorescence setup, which could be accurately correlated with the intraoperative light sources and fluorescence conditions, allowed for testing of new fluorescence visualisation techniques without having to overcome the obstacles of access to sterile neurosurgical theatres. The PpIX fluorescence, with its rapid photochemical reaction, was also characterised and a stable surrogate, using DCM laser dye, was proposed and validated. It is hoped that DCM will be used by other researchers as a stable alternative to PpIX. This setup was then used to test the hypothesis that the 5ALA endoscope was more sensitive at detecting visual fluorescence than the operative microscope and confirmed that this is indeed the case both in terms of visual detection and fluorescence spectrometry detection thresholds.

7.3.2 Future work

With the establishment of this laboratory setup our group are in the process of developing new methods of visualising 5ALA-induced fluorescence. Using the laboratory setup, preliminary testing of an ultra-low light level sensitive camera has been undertaken, and this has resulted in the awarding of funds to purchase such a camera to undertake further research using this technology.

The established setup will also be important in the development and validation an intraoperative fluorescence probe-based system which will be needed in order to carry out a head-to-head comparison of intraoperative fluorescence and Raman analysis to define glioma tumour boundaries

7.4 Comparing Raman spectroscopy and visual fluorescence

7.4.1 Conclusions

5ALA-induced fluorescence-guided surgery is an important tool used for the resection of high grade gliomas. The accuracy of any novel intraoperative Raman spectroscopy application used to define glioma tumour boundaries should therefore be compared with this current gold standard. Chapter 6 reported an experiment comparing the accuracy of Raman spectroscopy with 5ALA-induced visual fluorescence to predict the presence of infiltrating tumour from normal brain in fresh tissue samples. There are obvious

limitations to this study, as the Raman is *ex vivo* and the number of cases is small, but it does provide proof of concept that Raman has the potential to be much more sensitive in detecting tumour compared to visual fluorescence.

7.4.2 Future work

Future work should involve the development, and intraoperative assessment, of a Raman fibre optic probe-based system and a head-to-head comparison of this Raman system with a fluorescence probe. As outlined in Chapters 1 and 5, fluorescence probes have been shown to be able to identify fluorescence beyond the visual 5ALA-induced fluorescence boundary and thus it is important that the performance of an intraoperative Raman system to detect tumour boundaries is compared with that of an intraoperative fluorescence system. This also allows the notion of a multimodal classification model, using both fluorescence and Raman spectra, to be evaluated.

Some work has already been commenced along with our industry collaborator, Renishaw, to develop a probe-based Raman system. Several off-the-shelf Raman probes, compatible with the Renishaw Raman systems, have been considered as well as exploring the option of developing a purpose-built Raman probe. The ability of a probe to be sterilised and withstand the sterilisation process is an important consideration which has been partially explored. Rather than the routine high-temperature sterilisation processes used for most surgical instruments, many endoscopic instruments are sterilised using a low-temperature sterilisation system. The STERRAD® sterilisation system (Advances Sterilization Products, US)[302] is used in most UK hospitals for low-temperature sterilisation. Prior to sterilisation, all equipment is put through a disinfectant wash at 60 °C. To this end, a Raman probe (FlexiSpec Raman Probe, art photonics GmbH, Germany), compatible with the Renishaw RA800-series benchtop spectrometer, was tested. Testing was done to determine if the probe could be sterilised using the STERRAD® system and if it could withstand repeated sterilisations. After one round of disinfectant wash and STERRAD® sterilisation, the probe was tested by the sterilisation services quality control unit and was shown to be sterilised effectively and thus pass the quality control standard for use as a sterile surgical instrument. With regards optical integrity, the probe was tested by the manufacturers, art photonics GmbH, before and then after one, five and 20 cycles of disinfectant wash and STERRAD® sterilisation. Testing included evaluation of laser transmission, signal quality and background. There was no effect on these

parameters after one or five sterilisation cycles but, after 20 cycles, there was a significant reduction in both spectral quality and laser transmission. Microscopic inspection of the probe identified protrusion of epoxy glue onto the optical components as the most likely cause for this decrease in performance. These results have helped guide future probe development work and have led to the consideration of a disposable probe model where a probe is limited to a set number of surgeries and sterilisations before being discarded. Needless to say, this area requires much more research and development and the possibility of collaboration with a research group or company that already have a fully tested Raman probe system, sanctioned for intraoperative use, should also be considered.

The workflow design of an intraoperative study must involve recording visual fluorescence in a region of interest in the surgical resection cavity at the boundary of the tumour as defined by neuronavigation. The area should then be interrogated with both Raman and fluorescence probes. The area must then be accurately removed and *ex vivo* Raman and fluorescence probe measurements should be repeated before the sample is processed and undergoes full histopathological classification for the density of tumour within the sample. Care must be taken as to how the biopsy is practically performed, making sure it accurately correlates with the area interrogated with the probes. Given the importance of detecting the lowest level of tumour infiltration, it would be useful to test the Raman probe system on *ex vivo* tumour transition zones using cryosections in order to establish the threshold of detection which can be achieved by the system for detecting percentage tumour infiltration. It would also be useful to consult neuro-oncology surgeons to establish what a clinically useful tumour infiltration detection threshold would be. These parameters need to be established prior to undertaking any formal intraoperative trial in order to understand the degree of performance one is expecting from the system – if the surgeons feel they need to be able to detect tumour down to 5% tumour infiltration for a probe device to be useful, and the cryosection testing yields a threshold an order of magnitude greater than this, then further instrumentation development may be required before undertaking the *in vivo* study. On the other hand, as discussed in Chapter 1, the diffuse nature of these gliomas means that complete resection is never possible, and the surgeon is limited by functional boundaries where resection of further tissue could result in neurological disability for the patient. With is in mind, an intraoperative Raman or fluorescence probe only needs to perform better than the current gold standards of 5ALA-induced fluorescence-guided surgery in order to

show they have the potential to influence the extent of tumour resection and thus patient survival. It is also important that the false positive rate is accurately assessed to prevent the risk of resecting potentially normal brain, increasing the chance of neurological disability. The final stage in the evaluation of Raman and/or fluorescence intraoperative probe systems would therefore be a randomised trial comparing the use of the current gold standard techniques with and without the Raman and/or fluorescence probe technologies. The study outcomes should relate to patient survival and postoperative neurological disability.

7.5 Closing remarks

This thesis reports the first attempt to use Raman spectroscopy for the genetic sub-classification of gliomas, and is the largest study of its kind incorporating fresh tissue, cryosections, FFPE and glioma cell lines. The fluorescence work reported has developed a novel PpIX fluorescence standard and the fluorescence setup has been used to show that 5ALA-induced fluorescence is better visualised using the operative endoscope than the microscope. Finally, the first comparison of Raman spectroscopy with 5ALA-induced fluorescence is reported, demonstrating that Raman spectroscopy is significantly superior to visual fluorescence in predicting the presence of tumour at the resection margins. It is hoped that this body of work, after further development and validation, will result in Raman spectroscopy and fluorescence spectrometry applications that will facilitate intraoperative and clinical decision making in the treatment of patients with gliomas and, ultimately, improve the quality and length of life for these patients.

Appendix A: Coating protocol

Protocol for slide coating

1. Wash in 2.5% Decon for 5 mins
2. Wash in distilled water for 5 minutes
3. Rinse in 100% alcohol (repeat this twice with fresh alcohol)
4. Dry in the oven for 10 minutes
5. Place in 2% APTES (10 mL in 500 mL alcohol) for 5 minutes
6. Place in 100% alcohol for 5 minutes (repeat this twice with fresh alcohol)
7. Wash in distilled water for 5 minutes (repeat this twice with fresh alcohol)
8. Dry in oven for 10 minutes
9. Leave to cool in fume hood

Risk assessment

Chemical	COSHH Hazard	Risk Assessment
1. Alcohol (IMS)	Flammable	Avoid heat sources/flames. Refer to MRISK-ORH-NP 346
2. APES (3-AminoPropylTriEthoxy Silane)	Irritant/Corrosive	Handle with care in the fume hood. Wear gloves when handling. The working solution can be safely run to waste down the fume hood sink with plenty of running tap water

Appendix B: Fresh tissue

Two documents are included here. The first is the Fresh tissue for research – Standard Operating Procedure which was developed for this project. The second relates to the handling of fresh tissue for Raman analysis.

Fresh tissue for research – Standard Operating Procedure

This document outlines the standard procedure to be used when **fresh tissue samples** are to be taken for research.

1. Oxford Brain Bank (OBB) approval

1.1. Prior to the collection of any tissue the researcher needs to have gained approval from the OBB Access Committee by submission of an OBB Tissue Request form.

1.2. Once approval has been given an OBB reference number is issued by the OBB managers and a Material Transfer Agreement (MTA) for the specific project is formulated by the OBB managers with input from the researcher.

1.3. Every time fresh tissue is taken from a new patient a MTA must be completed.

2. Patient identification and consent

2.1. Appropriate patients are identified by the operating surgeon and researcher.

2.2. Operating surgeon completes NHS Consent Form 1 which includes the permission for use of surplus pathological tissue for research.

2.3. If patient gives consent for research then the researcher informs an OBB manager of the patient details.

2.4. Brain bank manager gives approval to collect fresh tissue from this patient under the OBB reference number*.

2.5. Researcher informs NHS lab of patient on day of surgery.

2.6. In theatre, the neuropathology request form is completed by the operating surgeon and ticked that consent for research has been given.

2.7. Researcher brings a copy of the Consent Form 1 to neuropathology at the start of the operation and this is scanned into Neuropathology system. The researcher must return the form to theatres.

Samples can only be used if the permission for use of surplus pathological tissue for research has been ticked on both Consent Form 1 and on the neuropathology request form.

3. Sample processing

3.1. Sample transferred from theatres to neuropathology using the current NHS standard protocol which includes appropriate sample labelling and signing out the sample from theatres.

3.2. Sample booked into NHS lab by lab technician as per standard lab protocols.

3.3. Lab technician checks copy of Consent Form 1 and consent box on neuropathology form and documents on Filemaker (including a note that Consent Form 1 has been copied).

3.4. Duty neuropathologist reviews sample and decides if enough tissue is available for diagnosis.

Fresh tissue can only be realised for research once the duty neuropathologist is happy that the NHS sample to be used for diagnostics is of sufficient quantity and quality. It is recommended that

best practice is for the first sample received by the lab is used solely for NHS diagnostics with only subsequent samples used for research.

3.5. Sample taken to cut up room and neuropathologist cuts it into 2 pieces:
  3.5.1. One piece returned to NHS lab technician for routine diagnostic analysis
  3.5.2. One piece given to researcher and labelled with SH number and section number for later correlation.
3.6. Researcher must make own risk assessment if any further sample processing is required in Neuropathology and this should be approved by the NHS lab manager.
3.7. Researcher must make own risk assessment for sample transportation out of Neuropathology.
3.8. Researcher must complete MTA form for that patient including:
  3.8.1. Confirmation that Consent Form 1 has been copied
  3.8.2. SH number
  3.8.3. Size and description of sample
3.9. MTA form given to OBB manager to file with OBB Tissue Request form
3.10. Lab technician must record on Filemaker worksheet that fresh tissue has been taken for research and from which sections
  3.10.1. Standard text: “Fresh tissue for research OBB <insert project number>”.

Fresh human tissue handling for Raman spectroscopy analysis

This document outlines the standard procedure for handling of fresh human tissue samples for Raman spectroscopy analysis.

Consent

Samples can only be used if the permission for use of pathological tissue has been ticked on both the patient consent form (Consent Form 1) and on the neuropathology request form.

Procedure

1. Samples booked into NHS lab under SH number.
2. Wearing appropriate PPE (lab coat, gloves, safety goggles) the sample allocated to research is collected in specimen pot from NHS laboratory after it has been booked in.
3. In cut-up room, sample is smeared into groove of pre-labelled a stainless steel slide.
4. Used surfaces in cut-up room cleaned with BioCleanse.
5. Slide placed in small white plastic clinical tray for transport to dark room.
6. New pair of gloves used to handle tray.
7. Tray with sample taken to dark room and slide placed in Raman microscope.
8. Microscope door closed and locked. Sign on microscope door: “Do not open – fresh tissue analysis in progress”.
9. Raman measurements carried out.
10. Slide taken back to cut-up room in clinical tray and cleaned with water and Biocleanse using paper towels, cotton buds and water. Waste disposed into orange clinical waste bin.
11. Removeable slide holder in Raman microscope cleaned with BioCleanse and water.
12. Used surfaces in dark room (should be minimal) cleaned with BioCleanse.

Risk assessment

HAZARDS	PERSONS AFFECTED	RISKS	CONTROL MEASURES
Exposure to fresh human tissue i.e. HIV, CJD	Any user sampling including but not limited to others working in the vicinity	Infection – VERY LOW	• Wear protective equipment including a laboratory coat, plastic apron, gloves, safety goggles, when in contact with fresh human tissue. • Decontaminate instruments and surfaces using the appropriate protocols • Dispose of all soiled paper towels, gloves, plastic aprons, etc into orange clinical bins and any sharps into sharps bin. • Any case with suspected or confirmed prion infection or any other category 3 pathogen will <u>not</u> be included in the study.
Contaminated instruments	N/A	Cross-contamination of instruments - LOW	Decontaminate instruments and surfaces using the appropriate protocols: BioCleanse for surfaces

COSHH

Substances used	Quantities used	Frequency of use	Hazards identified	Exposure route
5% Biocleanse™	About 20 mL	Approx 2-3 times a week	Risk of serious damage to eyes Harmful if swallowed Irritating to eyes and skin Causes burns	Skin, oral, inhalation.

Could a less hazardous substance (or form of the substance) be used instead? No

What measures have you taken to control risk?

PPE: Lab coat, solvent-resistant gloves, safety goggles.

Is health surveillance required? No	Training requirements: routine staff induction

Appendix C: Peak Allocation

a) Two group model and three group model analysis of IDH-mutant versus IDH-wildtype.

b) Three group model analysis of IDH-mutant versus oligodendroglioma (oligo)

The median wavenumber is shown in the first column and the model and tissue type shown across the top. For each model (Fresh, cryosection, FFPE, 2 group or 3 group etc.) the wavenumbers (wn) resolved from the -2ndD plot is listed with the corresponding p value (using Wilcoxon rank sum test) and the pathology group (path) with the most significant peak at that wavenumber indicated (m= IDH-mutant; w= IDH-wildtype; o= oligodendroglioma; e= equal in all pathologies). Green shading signifies a wavenumber with no significant difference ($p>0.01$). Yellow shading with red text signifies the 25% lowest p values for that model. To correlate the resolved peaks across models the number of tissue types/models with each peak is shown for the 2 group model, the LN18 cell models, the 3 group model IDH-mutant versus IDH-wildtype and a combination of all of these. Wavenumbers that are present in all models are shaded in blue. Similar correlation is done for the 3 group model analysis of IDH-mutant versus oligodendroglioma.

Wavenumber (cm-1)	2 group model																	3 group model (IDH-MUT minus IDH-WT)										Total No. of models with peak
	Fresh			Cryosection			FFPE			No. of tissue types with peak	LN18 Cryo			LN 18 FFPE			No. of LN18 cell preparation with peak	Fresh			Cryosection			FFPE			No. of tissue type with peak	
Median	wn	p value	Path	wn	p value	Path	wn	p value	Path		wn	p value	Path	wn	p value	Path		wn	p value	Path	wn	p value	Path	wn	p value	Path		
419	419	1.76E-01	m							1								419	7.44E-11	m							1	2
460	460	4.15E-07	w							1								460	2.87E-02	e	460	7.93E-28	w				2	3
491				491	8.22E-69	w				1	491	2.59E-01	w	491	4.57E-07	m	2				491	1.55E-125	w				1	4
517	517	3.31E-39	w							1								517	1.05E-14	w							1	2
529				529	3.93E-21	m	529	6.54E-53	m	2	529	1.76E-02	w	529	1.56E-08	m	2				529	4.70E-49	m	529	8.04E-75	m	2	6
540	540	1.24E-01	m							1								540	1.92E-05	e							1	2
565				566	5.88E-40	m	564	2.93E-20	m	2	566	4.00E-03	m				1				565	5.17E-163	m	565	1.52E-44	m	2	5
616	615	1.79E-81	m				617	4.58E-13	m	2				617	9.55E-51	m	1	615	9.59E-35	m	616	0.00E+00	m	616	6.69E-45	m	3	6
633	635	4.57E-64	m	633	0.00E+00	m	633	1.20E-08	m	3	633	1.82E-03	m	633	6.75E-37	m	2	633	3.42E-21	m	633	0.00E+00	m	633	2.64E-07	m	3	8
649	649	5.44E-32	m	649	0.00E+00	m	649	4.16E-09	m	3	649	7.81E-03	m	649	1.19E-33	m	2	649	3.17E-07	m	649	0.00E+00	m	649	1.53E-07	m	3	8
669				669	2.66E-01	w				1	670	1.16E-01	w				1				669	3.24E-27	m				1	3
710	710	1.80E-108	m							1								710	8.60E-64	m							1	2
724				723	4.96E-11	w	725	2.23E-01	m	2	724	1.05E-02	w	724	6.39E-01	e	2	725	3.71E-03	w	724	9.44E-01	m	724	1.74E-16	w	3	7
752	752	4.98E-01	w	752	1.19E-04	e	754	1.53E-25	w	3								752	2.70E-01	w	752	9.40E-06	m	752	2.33E-37	w	3	6
783				783	5.17E-210	w				1	783	6.96E-03	w	783	2.00E-95	m	2				783	7.52E-133	w				1	4
825	825	3.50E-36	m	825	8.20E-07	m	822	4.04E-03	m	3	825	1.58E-11	m	825	1.07E-04	m	2	825	5.24E-56	m	825	1.01E-53	w	825	6.76E-05	m	3	8
853	853	3.07E-199	w	855	4.90E-09	e	853	1.01E-108	w	3	853	1.90E-44	m	853	9.94E-01	e	2	853	5.50E-108	w	853	3.96E-175	w	853	6.36E-62	w	3	8
884	883	1.06E-18	m							1								885	1.31E-02	w							1	2
889				889	2.90E-18	w	889	9.50E-02	m	2	889	2.98E-05	m	889	6.34E-54	w	2				889	1.26E-71	w	889	1.97E-04	m	2	6
930				930	9.36E-29	w				1	930	8.72E-01	m				1										0	2
933	933	3.09E-108	w				932	1.61E-05	w	2								933	2.03E-44	w	933	3.61E-79	w	933	1.34E-11	w	3	5
963	963	7.51E-92	w				963	5.78E-63	m	2				963	1.28E-46	w	1	963	1.73E-51	w	963	7.00E-12	m	963	8.35E-01	m	3	6
1003	1003	0.00E+00	w	1003	4.97E-203	w	1003	1.17E-76	w	3	1003	1.54E-13	w	1003	1.17E-177	m	2	1003	0.00E+00	w	1003	1.22E-273	w	1003	1.08E-34	w	3	8
1036	1036	1.76E-68	m	1036	2.92E-05	e	1036	3.27E-59	w	3	1036	1.61E-05	m	1036	1.15E-29	m	2	1036	3.80E-12	m	1036	6.64E-07	w	1036	4.14E-07	w	3	8
1059	1059	0.00E+00	m							1								1059	0.00E+00	m							1	2
1087	1087	2.82E-167	m							1				1087	2.26E-101	w	1	1087	1.25E-109	m							1	3
1096				1098	1.01E-93	w	1094	9.69E-05	w	2	1095	5.02E-01	w	1095	1.26E-47	w	2				1096	1.19E-264	w	1096	7.32E-11	m	2	6
1124	1126	8.26E-01	m	1124	1.73E-01	e	1124	3.14E-02	e	3	1124	7.53E-07	m	1124	2.63E-13	w	2	1124	5.61E-15	m	1124	3.39E-33	w	1124	3.14E-39	m	3	8
1162	1160	1.51E-169	w				1164	6.28E-74	m	2	1163	2.05E-04	w				1	1160	1.53E-111	w	1162	1.85E-29	m	1162	6.15E-174	m	3	6
1206	1207	7.38E-239	w	1206	0.00E+00	w	1206	4.41E-02	w	3	1206	3.05E-01	w	1206	6.02E-93	m	2	1206	3.96E-147	w	1206	0.00E+00	w	1206	2.57E-15	m	3	8
1234				1233	1.89e-310	w	1235	2.70E-55	w	2	1234	8.45E-01	w	1234	4.74E-43	m	2				1233	0.00E+00	w	1233	3.42E-25	w	2	6
1265	1266	3.69E-118	w				1265	1.97E-41	m	2								1265	2.11E-121	w				1266	4.86E-09	m	2	4

a (continued)

Wavenumber (cm-1)	2 group model																3 group model (IDH-MUT minus IDH-WT)										Total No. of models with peak
	Fresh			Cryosection			FFPE			No. of tissue types with peak	LN18 Cryo			LN 18 FFPE			No. of LN18 cell preparation with peak	Fresh			Cryosection			FFPE			No. of tissue type with peak	
Median	wn	p value	Path	wn	p value	Path	wn	p value	Path		wn	p value	Path	wn	p value	Path		wn	p value	Path	wn	p value	Path	wn	p value	Path		
1305	1305	1.93E-01	e							1				1305	1.03E-82	w	1	1305	2.08E-28	m				1305	8.10E-133	m	2	4
1310				1311	5.79E-62	w	1309	8.61E-111	m	2	1310	1.29E-09	w				1				1310	8.59E-08	w				1	4
1337				1336	6.92E-08	w	1338	9.99E-14	m	2	1337	4.15E-01		1337	1.60E-90	m	2				1337	2.86E-13	m	1337	8.53E-01	e	2	6
1342	1342	3.41E-164	w							1								1342	3.67E-66	w							1	2
1445	1445	1.90E-89	m	1445	9.71E-03	m	1445	8.16E-42	m	3	1445	2.19E-03	w	1445	9.16E-156	w	2	1445	7.62E-60	m	1445	1.10E-89	w	1445	9.29E-129	m	3	8
1522	1522	8.27E-04	e							1								1522	1.56E-02	w	1522	0.00E+00	m				2	3
1553	1551	4.51E-29	m	1553	3.21E-135	m	1555	3.77E-25	w	3	1553	2.90E-04	w	1555	1.48E-13	m	2	1551	1.30E-04	m	1553	2.36E-198	m	1553	2.18E-69	w	3	8
1609				1609	1.32E-07	m	1609	7.17E-13	w	2	1609	3.31E-01	w	1609	6.77E-05	m	2				1609	3.58E-04	m	1609	6.78E-02	e	2	6
1611	1611	1.76E-51	m							1								1611	1.56E-40	w							1	2
1660	1660	3.50E-105	w	1660	8.13E-295		1660	6.32E-01	w	3	1660	4.70E-03	w	1660	5.35E-30	m	2	1660	1.86E-84	w	1660	0.00E+00	w	1660	3.62E-71	m	3	8
1740	1740	5.13E-102	m							1								1740	9.39E-96	m	1740	2.69E-31	m				2	3

b

Wavenumber (cm-1)	3 group model (IDH-MUT/IDH-WT versus Oligo)															No. of tissue types with peak
	Fresh					Cryosection					FFPE					
	wn	MUT vs Oligo		WT vs Oligo		wn	MUT vs Oligo		WT vs Oligo		wn	MUT vs Oligo		WT vs Oligo		
Median		p value	Path	p value	Path		p value	Path	p value	Path		p value	Path	p value	Path	
419	419	3.67E-40	m	5.07E-27	w											2
460	460	1.49E-12	m	6.36E-17	w	460	2.15E-33	o	5.49E-02	e						4
491						491	1.06E-83	o	4.40E-04	w						2
517	517	1.64E-35	m	1.86E-65	w											2
529						529	1.24E-32	m	3.68E-01	e	529	1.28E-04	e	1.76E-25	o	4
540	540	8.87E-15	m	7.77E-08	w											2
565						565	2.81E-227	m	8.83E-28	w	565	1.66E-10	m	1.53E-06	o	4
616	615	4.79E-67	o	2.51E-114	o	616	1.10E-159	m	2.24E-101	o	616	4.04E-01	o	8.93E-32	o	6
633	633	4.16E-75	o	4.55E-119	o	633	6.06E-12	m	0.00E+00	o	633	5.09E-64	o	5.14E-88	o	6
649	649	4.15E-60	o	1.08E-83	o	649	4.68E-142	m	6.10E-110	o	649	5.48E-59	o	9.86E-81	o	6
669						669	1.84E-178	m	2.85E-41	w						2
710	710	3.06E-35	o	1.20E-93	o											2
724	725	8.86E-07	o	4.95E-11	o	724	2.97E-55	m	4.30E-48	w	724	3.80E-148	o	4.13E-81	o	6
752	752	1.25E-01	e	5.11E-01	e	752	3.48E-97	m	8.17E-51	w	752	3.82E-108	o	9.33E-32	o	6
783						783	3.71E-50	o	2.02E-224	w						2
825	825	1.13E-30	m	5.34E-02	w	825	0.00E+00	o	4.74E-247	o	825	1.07E-21	m	3.19E-10	w	6
853	853	1.63E-44	m	1.21E-147	w	853	0.00E+00	o	1.62E-180	o	853	1.55E-71	m	1.38E-180	w	6
884	881	1.20E-157	o	9.58E-163	o											2
889	885	3.60E-148	o	1.71E-151	o											2
930						889	3.49E-117	o	8.29E-13	o	889	3.16E-34	m	8.14E-18	w	4
933	933	1.26E-73	m	5.35E-158	w	933	3.82E-71	o	6.83E-01	o	933	1.57E-07	o	6.17E-01	e	6
963	963	3.65E-27	m	1.31E-86	w	963	3.95E-27	m	2.08E-06	w	963	4.37E-106	o	8.80E-98	o	6
1003	1003	6.19E-24	m	2.46E-244	w	1003	3.63E-72	o	3.04E-50	w	1003	1.75E-36	m	7.79E-113	w	6
1036	1036	1.32E-139	o	1.16E-200	o	1036	1.87E-02	o	5.98E-02	o	1036	4.76E-212	m	3.81E-242	w	6
1059	1059	4.22E-75	o	1.21E-293	o											2
1087	1087	2.04E-21	o	1.07E-115	o											2
1096						1096	0.00E+00	o	2.75E-09	o	1096	0.00E+00	m	1.06E-304	w	4
1124	1124	2.38E-95	m	4.01E-70	w	1124	5.79E-131	o	1.03E-48	o	1124	3.53E-224	m	2.74E-79	w	6
1162	1160	9.87E-32	m	2.72E-116	w	1162	1.16E-258	m	2.91E-153	w	1162	3.44E-90	m	6.10E-01	e	6
1206	1206	7.21E-72	m	1.49E-186	w	1206	1.25E-06	o	1.75E-256	w	1206	1.22E-146	m	1.03E-71	w	6
1234						1233	5.11E-79	o	1.13E-108	w	1233	2.76E-14	m	1.83E-51	w	4
1265	1265	9.66E-11	o	3.83E-18	w						1266	5.79E-61	o	1.11E-94	o	4
1305	1305	7.48E-190	m	5.26E-168	w						1305	1.68E-16	m	9.12E-28	o	4
1310						1310	1.32E-133	m	5.50E-181	w						2
1337						1337	6.10E-181	m	2.77E-113	w	1337	1.65E-54	o	1.58E-56	o	4
1342	1342	3.09E-167	m	4.31E-243	w											2
1445	1445	1.47E-04	e	5.18E-60	o	1445	0.00E+00	o	3.45E-272	o	1445	5.43E-199	m	1.03E-16	w	6
1522	1522	2.32E-85	o	1.70E-44	o	1522	9.42E-215	m	9.91E-46	o						4
1553	1551	3.79E-83	o	3.95E-95	o	1553	1.85E-84	m	1.13E-24	o	1553	1.64E-131	o	2.71E-24	o	6
1609						1609	3.10E-03	o	2.02E-10	o	1609	9.36E-56	m	1.19E-53	w	4
1611	1611	4.27E-01	e	1.50E-23	w											2
1660	1660	1.47E-02	m	5.68E-44	w	1660	6.56E-82	o	1.52E-109	w	1660	0.00E+00	m	1.80E-189	w	6
1740	1740	5.99E-04	m	1.89E-24	o	1740	3.19E-39	o	9.19E-94	o						4

References

1. Reifenberger G, Wirsching HG, Knobbe-Thomsen CB, Weller M. Advances in the molecular genetics of gliomas - implications for classification and therapy. *Nat Rev Clin Oncol.* 2017; 14(7):434-452.

2. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. *New England Journal of Medicine.* 2005; 352(10):987-996.

3. Burnet NG, Jefferies SJ, Benson RJ, Hunt DP, Treasure FP. Years of life lost (YLL) from cancer is an important measure of population burden--and should be considered when allocating research funds. *Br J Cancer.* 2005; 92(2):241-245.

4. CRUK. "Beating Cancer Sooner" - Our Research Strategy: CRUK; 2014.

5. Louis DN, Perry A, Reifenberger G, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* 2016; 131(6):803-820.

6. Louis DN, Ohgak H, Wiestler OD, et al. *WHO classification of tumours of the central nervous system.* Revised 4th ed2016.

7. Goodenberger ML, Jenkins RB. Genetics of adult glioma. *Cancer Genetics.* 2012; 205(12):613-621.

8. Louis DN, Perry A, Burger P, et al. International Society Of Neuropathology--Haarlem consensus guidelines for nervous system tumor classification and grading. *Brain Pathol.* 2014; 24(5):429-435.

9. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science.* 2008; 321(5897):1807-1812.

10. Yan H, Parsons DW, Jin G, et al. IDH1 and IDH2 mutations in gliomas. *N Engl J Med.* 2009; 360(8):765-773.

11. Balss J, Meyer J, Mueller W, Korshunov A, Hartmann C, von Deimling A. Analysis of the IDH1 codon 132 mutation in brain tumors. *Acta Neuropathol.* 2008; 116(6):597-602.

12. Noushmehr H, Weisenberger DJ, Diefes K, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell.* 2010; 17(5):510-522.

13. Waitkus MS, Diplas BH, Yan H. Isocitrate dehydrogenase mutations in gliomas. *Neuro-Oncology.* 2016; 18(1):16-26.

14. Capper D, Zentgraf H, Balss J, Hartmann C, Deimling A. Monoclonal antibody specific for IDH1 R132H mutation. *Acta Neuropathologica.* 2009; 118(5):599-601.

15. Sanson M, Marie Y, Paris S, et al. Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *J Clin Oncol.* 2009; 27(25):4150-4154.

16. Weller M, Felsberg J, Hartmann C, et al. Molecular predictors of progression-free and overall survival in patients with newly diagnosed glioblastoma: a prospective translational study of the German Glioma Network. *J Clin Oncol.* 2009; 27(34):5743-5750.

17. van den Bent MJ, Dubbink HJ, Marie Y, et al. IDH1 and IDH2 mutations are prognostic but not predictive for outcome in anaplastic oligodendroglial tumors: a report of the European Organization for Research and Treatment of Cancer Brain Tumor Group. *Clin Cancer Res.* 2010; 16(5):1597-1604.

18. Sonoda Y, Kumabe T, Nakamura T, et al. Analysis of IDH1 and IDH2 mutations in Japanese glioma patients. *Cancer Sci.* 2009; 100(10):1996-1998.

19. Nobusawa S, Watanabe T, Kleihues P, Ohgaki H. IDH1 mutations as molecular signature and predictive factor of secondary glioblastomas. *Clin Cancer Res.* 2009; 15(19):6002-6007.

20. Gravendeel LA, Kouwenhoven MC, Gevaert O, et al. Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. *Cancer Res.* 2009; 69(23):9065-9072.

21. Wick W, Hartmann C, Engel C, et al. NOA-04 randomized phase III trial of sequential radiochemotherapy of anaplastic glioma with procarbazine, lomustine, and vincristine or temozolomide. *J Clin Oncol.* 2009; 27(35):5874-5880.

22. Dubbink HJ, Taal W, van Marion R, et al. IDH1 mutations in low-grade astrocytomas predict survival but not response to temozolomide. *Neurology.* 2009; 73(21):1792-1795.

23. Hartmann C, Hentschel B, Wick W, et al. Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. *Acta Neuropathol.* 2010; 120(6):707-718.
24. Houillier C, Wang X, Kaloshi G, et al. IDH1 or IDH2 mutations predict longer survival and response to temozolomide in low-grade gliomas. *Neurology.* 2010; 75(17):1560-1566.
25. SongTao Q, Lei Y, Si G, et al. IDH mutations predict longer survival and response to temozolomide in secondary glioblastoma. *Cancer Sci.* 2012; 103(2):269-273.
26. Tran AN, Lai A, Li S, et al. Increased sensitivity to radiochemotherapy in IDH1 mutant glioblastoma as demonstrated by serial quantitative MR volumetry. *Neuro Oncol.* 2014; 16(3):414-420.
27. Zetterling M, Berhane L, Alafuzoff I, Jakola AS, Smits A. Prognostic markers for survival in patients with oligodendroglial tumors; a single-institution review of 214 cases. *PLoS One.* 2017; 12(11):e0188419.
28. Cairncross G, Wang M, Shaw E, et al. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402. *J Clin Oncol.* 2013; 31(3):337-343.
29. van den Bent MJ, Brandes AA, Taphoorn MJ, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. *J Clin Oncol.* 2013; 31(3):344-350.
30. Buckner JC, Shaw EG, Pugh SL, et al. Radiation plus Procarbazine, CCNU, and Vincristine in Low-Grade Glioma. *N Engl J Med.* 2016; 374(14):1344-1355.
31. Baumert BG, Hegi ME, van den Bent MJ, et al. Temozolomide chemotherapy versus radiotherapy in high-risk low-grade glioma (EORTC 22033-26033): a randomised, open-label, phase 3 intergroup study. *Lancet Oncol.* 2016; 17(11):1521-1532.
32. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med.* 2005; 352(10):997-1003.
33. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 2009; 10(5):459-466.
34. Brain tumours (primary) and brain metastases in adults. Vol NG99: National Institute for Health and Care Excellence (NICE); 2018.
35. Watson LA, Goldberg H, Berube NG. Emerging roles of ATRX in cancer. *Epigenomics.* 2015; 7(8):1365-1378.
36. Brat DJ, Aldape K, Colman H, et al. cIMPACT-NOW update 3: recommended diagnostic criteria for "Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV". *Acta Neuropathol.* 2018; 136(5):805-810.
37. Ozawa M, Brennan PM, Zienius K, et al. Symptoms in primary care with time to diagnosis of brain tumours. *Family Practice.* 2018; 35(5):551-558.
38. Weller M, van den Bent M, Tonn JC, et al. European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. *The Lancet Oncology.* 2017; 18(6):e315-e329.
39. Excellence NIfHaC. National Institute for Health and Clinical Excellence. Guidance on Cancer

Services: Improving Outcomes for People With Brain and Other CNS Tumours: 2006; 2006.
40. Excellence NIfHaC. Guidance on Cancer Services. Improving Outcomes for People with Brain and Other CNS Tumours. In: NICE, ed. London2006.
41. Osswald M, Jung E, Sahm F, et al. Brain tumour cells interconnect to a functional and resistant network. *Nature.* 2015; 528(7580):93-98.
42. Sneed PK, Gutin PH, Larson DA, et al. Patterns of recurrence of glioblastoma multiforme after external irradiation followed by implant boost. *International Journal of Radiation Oncology*Biology*Physics.* 1994; 29(4):719-727.
43. Wallner KE, Galicich JH, Krol G, Arbit E, Malkin MG. Patterns of failure following treatment for glioblastoma multiforme and anaplastic astrocytoma. *Int J Radiat Oncol Biol Phys.* 1989; 16(6):1405-1409.

44. Gerin C, Pallud J, Deroulers C, et al. Quantitative characterization of the imaging limits of diffuse low-grade oligodendrogliomas. *Neuro-Oncology.* 2013; 15(10):1379-1388.
45. Höllerhage HG, Zumkeller M, Becker M, Dietz H. Influence of type and extent of surgery on early results and survival time in glioblastoma multiforme. *Acta Neurochirurgica.* 1991; 113(1-2):31-37.
46. Kowalczuk A, Macdonald RL, Amidei C, et al. Quantitative imaging study of extent of surgical resection and prognosis of malignant astrocytomas. *Neurosurgery.* 1997; 41(5):1028-1036; discussion 1036-1028.
47. Prados MD, Gutin PH, Phillips TL, et al. Highly anaplastic astrocytoma: A review of 357 patients treated between 1977 and 1989. *International Journal of Radiation Oncology*Biology*Physics.* 1992; 23(1):3-8.
48. McGirt MJ, Chaichana KL, Gathinji M, et al. Independent association of extent of resection with survival in patients with malignant brain astrocytoma. *J Neurosurg.* 2009; 110(1):156-162.
49. Barker FG, Prados MD, Chang SM, et al. Radiation response and survival time in patients with glioblastoma multiforme. 1996; 84(3):442.
50. Vecht CJ, Avezaat CJ, van Putten WL, Eijkenboom WM, Stefanko SZ. The influence of the extent of surgery on the neurological function and survival in malignant glioma. A retrospective analysis in 243 patients. *J Neurol Neurosurg Psychiatry.* 1990; 53(6):466-471.
51. Buckner JC, Schomberg PJ, McGinnis WL, et al. A phase III study of radiation therapy plus carmustine with or without recombinant interferon-α in the treatment of patients with newly diagnosed high-grade glioma. *Cancer.* 2001; 92(2):420-433.
52. Curran Jr WJ, Scott CB, Horton J, et al. Does extent of surgery influence outcome for astrocytoma with atypical or anaplastic foci (AAF)? A report from three Radiation Therapy Oncology Group (RTOG) trials. *Journal of Neuro-Oncology.* 1992; 12(3):219-227.
53. Dinapoli RP, Brown LD, Arusell RM, et al. Phase III comparative evaluation of PCNU and carmustine combined with radiation therapy for high-grade glioma. *Journal of Clinical Oncology.* 1993; 11(7):1316-1321.
54. Jeremic B, Grujicic D, Antunovic V, Djuric L, Stojanovic M, Shibamoto Y. Influence of extent of surgery and tumor location on treatment outcome of patients with glioblastoma multiforme treated with combined modality approach. *Journal of Neuro-Oncology.* 1994; 21(2):177-185.
55. Lamborn KR, Prados MD, Chang SM. Prognostic factors for survival of patients with glioblastoma: Recursive partitioning analysis. *Neuro-Oncology.* 2004; 6(3):227-235.
56. Nitta T, Sato K. Prognostic implications of the extent of surgical resection in patients with intracranial malignant gliomas. *Cancer.* 1995; 75(11):2727-2731.
57. Nomiya T, Nemoto K, Kumabe T, Takai Y, Yamada S. Prognostic significance of surgery and radiation therapy in cases of anaplastic astrocytoma: retrospective analysis of 170 cases. 2007; 106(4):575.
58. Oszvald Á, Güresir E, Setzer M, et al. Glioblastoma therapy in the elderly and the importance of the extent of resection regardless of age. 2012; 116(2):357.
59. Shibamoto Y, Yamashita J, Takahashi M, Yamasaki T, Kikuchi H, Abe M. Supratentorial malignant glioma: an analysis of radiation therapy in 178 cases. *Radiotherapy and Oncology.* 1990; 18(1):9-17.
60. Simpson JR, Horton J, Scott C, et al. Influence of location and extent of surgical resection on survival of patients with glioblastoma multiforme: Results of three consecutive radiation therapy oncology group (RTOG) clinical trials. *International Journal of Radiation Oncology*Biology*Physics.* 1993; 26(2):239-244.
61. Stark AM, Nabavi A, Mehdorn HM, Blömer U. Glioblastoma multiforme—report of 267 cases treated at a single institution. *Surgical Neurology.* 2005; 63(2):162-169.
62. Ushio Y, Kochi M, Hamada JI, Kai Y, Nakamura H. Effect of surgical removal on survival and quality of life in patients with supratentorial glioblastoma. *Neurologia Medico-Chirurgica.* 2005; 45(9):454-460.
63. Keles GE, Anderson B, Berger MS. The effect of extent of resection on time to tumor progression and survival in patients with glioblastoma multiforme of the cerebral hemisphere. *Surg Neurol.* 1999; 52(4):371-379.

64. G. Evren Keles, Edward F. Chang, Kathleen R. Lamborn, et al. Volumetric extent of resection and residual contrast enhancement on initial surgery as predictors of outcome in adult patients with hemispheric anaplastic astrocytoma. *Journal of Neurosurgery.* 2006; 105(1):34-40.

65. Lacroix M, Abi-Said D, Fourney DR, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. *J Neurosurg.* 2001; 95(2):190-198.

66. Mark E. Oppenlander, Andrew B. Wolf, Laura A. Snyder, et al. An extent of resection threshold for recurrent glioblastoma and its risk for neurological morbidity. *Journal of Neurosurgery.* 2014; 120(4):846-853.

67. Orringer D, Lau D, Khatri S, et al. Extent of resection in patients with glioblastoma: limiting factors, perception of resectability, and effect on survival. *J Neurosurg.* 2012; 117(5):851-859.

68. Sanai N, Polley MY, McDermott MW, Parsa AT, Berger MS. An extent of resection threshold for newly diagnosed glioblastomas. *J Neurosurg.* 2011; 115(1):3-8.

69. Pope WB, Sayre J, Perlina A, Villablanca JP, Mischel PS, Cloughesy TF. MR imaging correlates of survival in patients with high-grade gliomas. *American Journal of Neuroradiology.* 2005; 26(10):2466-2474.

70. Sanai N, Berger MS. Glioma extent of resection and its impact on patient outcome. *Neurosurgery.* 2008; 62(4):753-764; discussion 264-756.

71. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? *Journal of Neurosurgery.* 2016; 124(4):977-988.

72. Chaichana KL, Jusue-Torres I, Navarro-Ramirez R, et al. Establishing percent resection and residual volume thresholds affecting survival and recurrence for patients with newly diagnosed intracranial glioblastoma. *Neuro-Oncology.* 2013; 16(1):113-122.

73. Stummer W, Pichlmeier U, Meinel T, et al. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *Lancet Oncol.* 2006; 7(5):392-401.

74. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. *The Lancet Oncology.* 2011; 12(11):997-1003.

75. Brown TJ, Brennan MC, Li M, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. *JAMA Oncol.* 2016; 2(11):1460-1469.

76. Aghi MK, Nahed BV, Sloan AE, Ryken TC, Kalkanis SN, Olson JJ. The role of surgery in the management of patients with diffuse low grade glioma: A systematic review and evidence-based clinical practice guideline. *J Neurooncol.* 2015; 125(3):503-530.

77. Ahmadi R, Rezvan A, Dictus C, et al. Long-term outcome and survival of surgically treated supratentorial low-grade glioma in adult patients. *Acta Neurochir (Wien).* 2009; 151(11):1359-1365.

78. Karim AB, Maat B, Hatlevoll R, et al. A randomized trial on dose-response in radiation therapy of low-grade cerebral glioma: European Organization for Research and Treatment of Cancer (EORTC) Study 22844. *Int J Radiat Oncol Biol Phys.* 1996; 36(3):549-556.

79. Leighton C, Fisher B, Bauman G, et al. Supratentorial low-grade glioma in adults: an analysis of prognostic factors and timing of radiation. *Journal of Clinical Oncology.* 1997; 15(4):1294-1301.

80. Lote K, Egeland T, Hager B, et al. Survival, prognostic factors, and therapeutic efficacy in low-grade glioma: a retrospective study in 379 patients. *Journal of Clinical Oncology.* 1997; 15(9):3129-3140.

81. Nicolato A, Gerosa MA, Fina P, Iuzzolino P, Giorgiutti F, Bricolo A. Prognostic factors in low-grade supratentorial astrocytomas: A uni-multivariate statistical analysis in 76 surgically treated adult patients. *Surgical Neurology.* 1995; 44(3):208-223.

82. Scerrati M, Roselli R, Iacoangeli M, Pompucci A, Rossi GF. Prognostic factors in low grade (WHO grade II) gliomas of the cerebral hemispheres: the role of surgery. *Journal of Neurology, Neurosurgery & Psychiatry.* 1996; 61(3):291-296.

83. Ito S, Chandler KL, Prados MD, et al. Proliferative potential and prognostic evaluation of low-grade astrocytomas. *Journal of Neuro-Oncology.* 1994; 19(1):1-9.

84. Peraud A, Ansari H, Bise K, Reulen H-J. Clinical Outcome of Supratentorial Astrocytoma WHO Grade II. *Acta Neurochirurgica.* 1998; 140(12):1213-1222.
85. Rajan B, Pickuth D, Ashley S, et al. The management of histologically unverified presumed cerebral gliomas with radiotherapy. *International Journal of Radiation Oncology*Biology*Physics.* 1994; 28(2):405-413.
86. Shaw E, Arusell R, Scheithauer B, et al. Prospective Randomized Trial of Low- Versus High-Dose Radiation Therapy in Adults With Supratentorial Low-Grade Glioma: Initial Report of a North Central Cancer Treatment Group/Radiation Therapy Oncology Group/Eastern Cooperative Oncology Group Study. *Journal of Clinical Oncology.* 2002; 20(9):2267-2276.
87. Nakamura M, Konishi N, Tsunoda S, et al. Analysis of Prognostic and Survival Factors Related to Treatment of Low-Grade Astrocytomas in Adults. *Oncology.* 2000; 58(2):108-116.
88. Yeh S-A, Ho J-T, Lui C-C, Huang Y-J, Hsiung C-Y, Huang E-Y. Treatment outcomes and prognostic factors in patients with supratentorial low-grade gliomas. *The British Journal of Radiology.* 2005; 78(927):230-235.
89. Jakola AS, Myrmel KS, Kloster R, et al. Comparison of a Strategy Favoring Early Surgical Resection vs a Strategy Favoring Watchful Waiting in Low-Grade GliomasSurgical Resection vs Waiting in Low-Grade Gliomas. *JAMA.* 2012; 308(18):1881-1888.
90. Philippon JH, Clemenceau SH, Fauchon FH, Foncin JF. Supratentorial Low-Grade Astrocytomas in Adults. *Neurosurgery.* 1993; 32(4):554-559.
91. North CA, North RB, Epstein JA, Piantadosi S, Wharam MD. Low-grade cerebral astrocytomas: Survival and quality of life after radiation therapy. *Cancer.* 1990; 66(1):6-14.
92. Claus EB, Horlacher A, Hsu L, et al. Survival rates in patients with low-grade glioma after intraoperative magnetic resonance image guidance. *Cancer.* 2005; 103(6):1227-1233.
93. Incekara F, Olubiyi O, Ozdemir A, Lee T, Rigolo L, Golby A. The Value of Pre- and Intraoperative Adjuncts on the Extent of Resection of Hemispheric Low-Grade Gliomas: A Retrospective Analysis. *J Neurol Surg A Cent Eur Neurosurg.* 2016; 77(02):079-087.
94. Smith JS, Chang EF, Lamborn KR, et al. Role of extent of resection in the long-term outcome of low-grade hemispheric gliomas. *J Clin Oncol.* 2008; 26(8):1338-1345.
95. Snyder LA, Wolf AB, Oppenlander ME, et al. The impact of extent of resection on malignant transformation of pure oligodendrogliomas. 2014; 120(2):309.
96. Duffau H. Long-term outcomes after supratotal resection of diffuse low-grade gliomas: a consecutive series with 11-year follow-up. *Acta Neurochirurgica.* 2016; 158(1):51-58.
97. Beiko J, Suki D, Hess KR, et al. IDH1 mutant malignant astrocytomas are more amenable to surgical resection and have a survival benefit associated with maximal surgical resection. *Neuro Oncol.* 2014; 16(1):81-91.
98. Kawaguchi T, Sonoda Y, Shibahara I, et al. Impact of gross total resection in patients with WHO grade III glioma harboring the IDH 1/2 mutation without the 1p/19q co-deletion. *J Neurooncol.* 2016; 129(3):505-514.
99. Cordier D, Goze C, Schadelin S, Rigau V, Mariani L, Duffau H. A better surgical resectability of WHO grade II gliomas is independent of favorable molecular markers. *J Neurooncol.* 2015; 121(1):185-193.
100. Wijnenga MMJ, French PJ, Dubbink HJ, et al. The impact of surgery in molecularly defined low-grade glioma: an integrated clinical, radiological, and molecular analysis. *Neuro Oncol.* 2018; 20(1):103-112.
101. Yordanova YN, Duffau H. Supratotal resection of diffuse gliomas - an overview of its multifaceted implications. *Neurochirurgie.* 2017; 63(3):243-249.
102. Duffau H. Higher-Order Surgical Questions for Diffuse Low-Grade Gliomas: Supramaximal Resection, Neuroplasticity, and Screening. *Neurosurgery Clinics of North America.* 2019; 30(1):119-128.
103. Esquenazi Y, Friedman E, Liu Z, Zhu J-J, Hsu S, Tandon N. The Survival Advantage of "Supratotal" Resection of Glioblastoma Using Selective Cortical Mapping and the Subpial technique. *Neurosurgery.* 2017.
104. Livermore LJ, Williams S, Clifton L, et al. FUNCTIONALLY GUIDED SUPRAMAXIMAL RESECTION OF IDH-WILDTYPE GLIOBLASTOMAS AND THE EFFECT ON PROGRESSION FREE SURVIVAL. *Neuro-Oncology.* 2018; 20(suppl_5):v346-v347.

105. Gulati S, Jakola AS, Nerland US, Weber C, Solheim O. The Risk of Getting Worse: Surgically Acquired Deficits, Perioperative Complications, and Functional Outcomes After Primary Resection of Glioblastoma. *World Neurosurgery.* 2011; 76(6):572-579.
106. Jenkinson MD, Barone DG, Bryant A, et al. Intraoperative imaging technology to maximise extent of resection for glioma. *Cochrane Database Syst Rev.* 2018; 1:CD012788.
107. Schneider JP, Trantakis C, Rubach M, et al. Intraoperative MRI to guide the resection of primary supratentorial glioblastoma multiforme—a quantitative radiological analysis. *Neuroradiology.* 2005; 47(7):489-500.
108. Rao G, Weinberg JS, Hatiboglu MA, et al. IMPACT OF INTRAOPERATIVE HIGH-FIELD MAGNETIC RESONANCE IMAGING GUIDANCE ON GLIOMA SURGERY: A PROSPECTIVE VOLUMETRIC ANALYSIS. *Neurosurgery.* 2009; 64(6):1073-1081.
109. Hermann E, Franz K, Senft C, Gasser T, Seifert V. Usefulness of Intraoperative Ultra Low-field Magnetic Resonance Imaging in Glioma Surgery. *Operative Neurosurgery.* 2008; 63(suppl_4):ONS257-ONS267.
110. Senft C, Franz K, Ulrich CT, et al. Low field intraoperative MRI-guided surgery of gliomas: A single center experience. *Clinical Neurology and Neurosurgery.* 2010; 112(3):237-243.
111. Staubert A, Wirtz CR, Bonsanto MM, et al. Clinical Evaluation and Follow-up Results for Intraoperative Magnetic Resonance Imaging in Neurosurgery. *Neurosurgery.* 2000; 46(5):1112-1122.
112. Moiyadi AV, Shetty PM, Mahajan A, Udare A, Sridhar E. Usefulness of three-dimensional navigable intraoperative ultrasound in resection of brain tumors with a special emphasis on malignant gliomas. *Acta Neurochirurgica.* 2013; 155(12):2217-2225.
113. Solheim O, Selbekk T, Jakola AS, Unsgård G. Ultrasound-guided operations in unselected high-grade gliomas—overall results, impact of image quality and patient selection. *Acta Neurochirurgica.* 2010; 152(11):1873-1886.
114. Gerganov VM, Samii A, Giordano M, Samii M, Fahlbusch R. Two-dimensional high-end ultrasound imaging compared to intraoperative MRI during resection of low-grade gliomas. *Journal of Clinical Neuroscience.* 2011; 18(5):669-673.
115. Shinoda J, Yano H, Yoshimura S, et al. Fluorescence-guided resection of glioblastoma multiforme by using high-dose fluorescein sodium. Technical note. *Journal of neurosurgery.* 2003; 99(3):597-603.
116. Murray KJ. Improved surgical resection of human brain tumors: Part I. A preliminary study. *Surgical neurology.* 1982; 17(5):316-319.
117. Li Y, Rey-Dios R, Roberts DW, Valdes PA, Cohen-Gadol AA. Intraoperative Fluorescence-Guided Resection of High-Grade Gliomas: A Comparison of the Present Techniques and Evolution of Future Strategies. *World neurosurgery.* 2013.
118. Koc K, Anik I, Cabuk B, Ceylan S. Fluorescein sodium-guided surgery in glioblastoma multiforme: a prospective evaluation. *Br J Neurosurg.* 2008; 22(1):99-103.
119. Liu JG, Yang SF, Liu YH, Wang X, Mao Q. Magnetic resonance diffusion tensor imaging with fluorescein sodium dyeing for surgery of gliomas in brain motor functional areas. *Chin Med J (Engl).* 2013; 126(13):2418-2423.
120. Li Y, Rey-Dios R, Roberts DW, Valdés PA, Cohen-Gadol AA. Intraoperative Fluorescence-Guided Resection of High-Grade Gliomas: A Comparison of the Present Techniques and Evolution of Future Strategies. *World Neurosurgery.* 2014; 82(1):175-185.
121. Stummer W, Suero Molina E. Fluorescence Imaging/Agents in Tumor Resection. *Neurosurg Clin N Am.* 2017; 28(4):569-583.
122. Stummer W. Response to Journal Club: 5-Aminolevulinic Acid–Derived Tumor Fluorescence: The Diagnostic Accuracy of Visible Fluorescence Qualities as Corroborated by Spectrometry and Histology and Postoperative Imaging. *Neurosurgery.* 2015; 76(2):230-231.
123. Colditz MJ, Leyen K, Jeffree RL. Aminolevulinic acid (ALA)-protoporphyrin IX fluorescence guided tumour resection. Part 2: theoretical, biochemical and practical aspects. *J Clin Neurosci.* 2012; 19(12):1611-1616.
124. Stummer W, Stocker S, Novotny A, et al. In vitro and in vivo porphyrin accumulation by C6 glioma cells after exposure to 5-aminolevulinic acid. *J Photochem Photobiol B.* 1998; 45(2-3):160-169.

125. Friesen SA, Hjortland GO, Madsen SJ, et al. 5-Aminolevulinic acid-based photodynamic detection and therapy of brain tumors (review). *Int J Oncol.* 2002; 21(3):577-582.
126. Collaud S, Juzeniene A, Moan J, Lange N. On the selectivity of 5-aminolevulinic acid-induced protoporphyrin IX formation. *Curr Med Chem Anticancer Agents.* 2004; 4(3):301-316.
127. Peng Q, Warloe T, Berg K, et al. 5-Aminolevulinic acid-based photodynamic therapy. Clinical research and future challenges. *Cancer.* 1997; 79(12):2282-2308.
128. Coburger J, Engelke J, Scheuerle A, et al. Tumor detection with 5-aminolevulinic acid fluorescence and Gd-DTPA–enhanced intraoperative MRI at the border of contrast-enhancing lesions: a prospective study based on histopathological assessment. 2014; 36(2):E3.
129. Stummer W, Novotny A, Stepp H, Goetz C, Bise K, Reulen HJ. Fluorescence-guided resection of glioblastoma multiforme utilizing 5-ALA-induced porphyrins: a prospective study in 52 consecutive patients. 2000; 93(6):1003.
130. Idoate MA, Díez Valle R, Echeveste J, Tejada S. Pathological characterization of the glioblastoma border as shown during surgery using 5-aminolevulinic acid-induced fluorescence. *Neuropathology.* 2011; 31(6):575-582.
131. Roberts DW, Valdes PA, Harris BT, et al. Coregistered fluorescence-enhanced tumor resection of malignant glioma: relationships between delta-aminolevulinic acid-induced protoporphyrin IX fluorescence, magnetic resonance imaging enhancement, and neuropathological parameters. Clinical article. *J Neurosurg.* 2011; 114(3):595-603.
132. Yamada S, Muragaki Y, Maruyama T, Komori T, Okada Y. Role of neurochemical navigation with 5-aminolevulinic acid during intraoperative MRI-guided resection of intracranial malignant gliomas. *Clin Neurol Neurosurg.* 2015; 130:134-139.
133. Eljamel S. 5-ALA Fluorescence Image Guided Resection of Glioblastoma Multiforme: A Meta-Analysis of the Literature. *Int J Mol Sci.* 2015; 16(5):10443-10456.
134. Stummer W, Tonn JC, Goetz C, et al. 5-Aminolevulinic acid-derived tumor fluorescence: the diagnostic accuracy of visible fluorescence qualities as corroborated by spectrometry and histology and postoperative imaging. *Neurosurgery.* 2014; 74(3):310-319; discussion 319-320.
135. Valdes PA, Leblond F, Kim A, et al. Quantitative fluorescence in intracranial tumor: implications for ALA-induced PpIX as an intraoperative biomarker. *J Neurosurg.* 2011; 115(1):11-17.
136. Puppa AD, Rustemi O, Gioffrè G, et al. Predictive value of intraoperative 5-aminolevulinic acid–induced fluorescence for detecting bone invasion in meningioma surgery. 2014; 120(4):840.
137. Potapov AA, Goryaynov SA, Okhlopkov VA, et al. Laser biospectroscopy and 5-ALA fluorescence navigation as a helpful tool in the meningioma resection. *Neurosurgical Review.* 2016; 39(3):437-447.
138. Motekallemi A, Jeltema HR, Metzemaekers JD, van Dam GM, Crane LM, Groen RJ. The current status of 5-ALA fluorescence-guided resection of intracranial meningiomas-a critical review. *Neurosurg Rev.* 2015; 38(4):619-628.
139. Aldave G, Tejada S, Pay E, et al. Prognostic value of residual fluorescent tissue in glioblastoma patients after gross total resection in 5-aminolevulinic Acid-guided surgery. *Neurosurgery.* 2013; 72(6):915-920; discussion 920-911.
140. Schucht P, Knittel S, Slotboom J, et al. 5-ALA complete resections go beyond MR contrast enhancement: shift corrected volumetric analysis of the extent of resection in surgery for glioblastoma. *Acta Neurochirurgica.* 2014; 156(2):305-312.
141. Stummer W, Reulen HJ, Meinel T, et al. Extent of resection and survival in glioblastoma multiforme: identification of and adjustment for bias. *Neurosurgery.* 2008; 62(3):564-576; discussion 564-576.
142. Díez Valle R, Tejada Solis S, Idoate Gastearena MA, García de Eulate R, Domínguez Echávarri P, Aristu Mendiroz J. Surgery guided by 5-aminolevulinic fluorescence in glioblastoma: volumetric analysis of extent of resection in single-center experience. *Journal of Neuro-Oncology.* 2011; 102(1):105-113.
143. Schucht P, Beck J, Abu-Isa J, et al. Gross total resection rates in contemporary glioblastoma surgery: Results of an institutional protocol combining 5-aminolevulinic acid intraoperative fluorescence imaging and brain mapping. *Neurosurgery.* 2012; 71(5):927-935.

144. Della Puppa A, De Pellegrin S, d'Avella E, et al. 5-aminolevulinic acid (5-ALA) fluorescence guided surgery of high-grade gliomas in eloquent areas assisted by functional mapping. Our experience and review of the literature. *Acta Neurochirurgica*. 2013; 155(6):965-972.
145. Coburger J, Hagel V, Wirtz CR, König R. Surgery for Glioblastoma: Impact of the Combined Use of 5-Aminolevulinic Acid and Intraoperative MRI on Extent of Resection and Survival. *PLOS ONE*. 2015; 10(6):e0131872.
146. Stummer W, Stepp H, Wiestler OD, Pichlmeier U. Randomized, Prospective Double-Blinded Study Comparing 3 Different Doses of 5-Aminolevulinic Acid for Fluorescence-Guided Resections of Malignant Gliomas. *Neurosurgery*. 2017; 81(2):230-239.
147. Håberg AK, Berntsen EM, Selbekk T, et al. Functional Magnetic Resonance Imaging and Diffusion Tensor Tractography Incorporated Into an Intraoperative 3-Dimensional Ultrasound-Based Neuronavigation System: Impact on Therapeutic Strategies, Extent of Resection, and Clinical Outcome. *Neurosurgery*. 2010; 67(2):251-264.
148. Romano A, Ferrante M, Cipriani V, et al. Role of magnetic resonance tractography in the preoperative planning and intraoperative assessment of patients with intra-axial brain tumours. *La radiologia medica*. 2007; 112(6):906-920.
149. Romano A, D'Andrea G, Minniti G, et al. Pre-surgical planning and MR-tractography utility in brain tumour resection. *European Radiology*. 2009; 19(12):2798.
150. Yordanka N. Yordanova, Sylvie Moritz-Gasser, Hugues Duffau. Awake surgery for WHO Grade II gliomas within "noneloquent" areas in the left dominant hemisphere: toward a "supratotal" resection. *Journal of Neurosurgery*. 2011; 115(2):232-239.
151. Duffau H, Gatignol P, Mandonnet E, Capelle L, Taillandier L. Intraoperative subcortical stimulation mapping of language pathways in a consecutive series of 115 patients with Grade II glioma in the left dominant hemisphere. 2008; 109(3):461.
152. Duffau H. A personal consecutive series of surgically treated 51 cases of insular WHO Grade II glioma: advances and limitations. 2009; 110(4):696.
153. De Witt Hamer PC, Robles SG, Zwinderman AH, Duffau H, Berger MS. Impact of intraoperative stimulation brain mapping on glioma surgery outcome: a meta-analysis. *J Clin Oncol*. 2012; 30(20):2559-2565.
154. Keles GE, Lundin DA, Lamborn KR, Chang EF, Ojemann G, Berger MS. Intraoperative subcortical stimulation mapping for hemispheric perirolandic gliomas located within or adjacent to the descending motor pathways: evaluation of morbidity and assessment of functional outcome in 294 patients. 2004; 100(3):369.
155. Eisner W, Burtscher J, Bale R, et al. Use of neuronavigation and electrophysiology in surgery of subcortically located lesions in the sensorimotor strip. *Journal of Neurology, Neurosurgery & Psychiatry*. 2002; 72(3):378-381.
156. Wu JS, Zhou LF, Tang WJ, et al. Clinical evaluation and follow-up outcome of diffusion tensor imaging-based functional neuronavigation: a prospective, controlled study in patients with gliomas involving pyramidal tracts. *Neurosurgery*. 2007; 61(5):935-948; discussion 948-939.
157. Kumar A, Chandra PS, Sharma BS, et al. The role of neuronavigation-guided functional MRI and diffusion tensor tractography along with cortical stimulation in patients with eloquent cortex lesions. *British Journal of Neurosurgery*. 2014; 28(2):226-233.
158. Bello L, Castellano A, Fava E, et al. Intraoperative use of diffusion tensor imaging fiber tractography and subcortical mapping for resection of gliomas: technical considerations. *Neurosurg Focus*. 2010; 28(2):E6.
159. Mikuni N, Okada T, Enatsu R, et al. Clinical impact of integrated functional neuronavigation and subcortical electrical stimulation to preserve motor function during resection of brain tumors. 2007; 106(4):593.
160. Ma R, Livermore LJ, Plaha P. Fast Track Recovery Program After Endoscopic and Awake Intraparenchymal Brain Tumor Surgery. *World Neurosurg*. 2016; 93:246-252.
161. Raman CV. A new radiation. *Indian J Physcis*. 1928(2):387.
162. *Nobel Lectures, Physics 1922-1941*. Amsterdam: Elsevier Publishing Company; 1965.
163. Raman CV, Krishnan KS. A New Type of Secondary Radiation. *Nature*. 1928; 121:501.
164. Butler HJ, Ashton L, Bird B, et al. Using Raman spectroscopy to characterize biological materials. *Nat Protoc*. 2016; 11(4):664-687.

165. Chase B. A New Generation of Raman Instrumentation. *Applied Spectroscopy.* 1994; 48(7):14A-19A.
166. Hollon T, Lewis S, Freudiger CW, Xie XS, Orringer DA. Improving the accuracy of brain tumor surgery via Raman-based technology. *Neurosurgical Focus.* 2016; 40(3):E9.
167. Broadbent B, Tseng J, Kast R, et al. Shining light on neurosurgery diagnostics using Raman spectroscopy. *J Neurooncol.* 2016; 130(1):1-9.
168. Mizuno A, Hayashi T, Tashibu K, Maraishi S, Kawauchi K, Ozaki Y. Near-infrared FT-Raman spectra of the rat brain tissues. *Neuroscience Letters.* 1992; 141(1):47-52.
169. Mizuno A, Kitajima H, Kawauchi K, Muraishi S, Ozaki Y. Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours. *Journal of Raman Spectroscopy.* 1994; 25(1):25-29.
170. Krafft C, Steiner G, Beleites C, Salzer R. Disease recognition by infrared and Raman spectroscopy. *Journal of Biophotonics.* 2009; 2(1-2):13-28.
171. Tanahashi K, Natsume A, Ohka F, et al. Assessment of tumor cells in a mouse model of diffuse infiltrative glioma by Raman spectroscopy. *BioMed Research International.* 2014; 2014:860241.
172. Amharref N, Beljebbar A, Dukic S, et al. Discriminating healthy from tumor and necrosis tissue in rat brain tissue samples by Raman spectral imaging. *Biochimica et Biophysica Acta.* 2007; 1768(10):2605-2615.
173. Gajjar K, Heppenstall LD, Pang W, et al. Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis. *Anal Methods.* 2012; 5:89-102.
174. Krafft C, Neudert L, Simat T, Salzer R. Near infrared Raman spectra of human brain lipids. *Spectrochimica Acta. Part A, Molecular & Biomolecular Spectroscopy.* 2005; 61(7):1529-1535.
175. Kohler M, Machill S, Salzer R, Krafft C. Characterization of lipid extracts from brain tissue and tumors using Raman spectroscopy and mass spectrometry. *Analytical & Bioanalytical Chemistry.* 2009; 393(5):1513-1520.
176. Krafft C, Sobottka SB, Schackert G, Salzer R. Near infrared Raman spectroscopic mapping of native brain tissue and intracranial tumors. *Analyst.* 2005; 130(7):1070-1077.
177. Krafft C, Belay B, Bergner N, et al. Advances in optical biopsy--correlation of malignancy and cell density of primary brain tumors using Raman microspectroscopic imaging. *Analyst.* 2012; 137(23):5533-5537.
178. Bergner N, Krafft C, Geiger KD, Kirsch M, Schackert G, Popp J. Unsupervised unmixing of Raman microspectroscopic images for morphochemical analysis of non-dried brain tumor specimens. *Analytical & Bioanalytical Chemistry.* 2012; 403(3):719-725.
179. Bergner N, Medyukhina A, Geiger KD, et al. Hyperspectral unmixing of Raman micro-images for assessment of morphological and chemical parameters in non-dried brain tumor specimens. *Analytical & Bioanalytical Chemistry.* 2013; 405(27):8719-8728.
180. Leslie DG, Kast RE, Poulik JM, et al. Identification of pediatric brain neoplasms using Raman spectroscopy. *Pediatric Neurosurgery.* 2012; 48(2):109-117.
181. Auner AW, Kast RE, Rabah R, Poulik JM, Klein MD. Conclusions and data analysis: a 6-year study of Raman spectroscopy of solid tumors at a major pediatric institute. *Pediatric Surgery International.* 2013; 29(2):129-140.
182. Kalkanis SN, Kast RE, Rosenblum ML, et al. Raman spectroscopy to distinguish grey matter, necrosis, and glioblastoma multiforme in frozen tissue sections. *Journal of Neuro-Oncology.* 2014; 116(3):477-485.
183. Kast RE, Auner GW, Rosenblum ML, et al. Raman molecular imaging of brain frozen tissue sections. *J Neurooncol.* 2014.
184. Kast R, Auner G, Yurgelevic S, et al. Identification of regions of normal grey matter and white matter from pathologic glioblastoma and necrosis in frozen sections using Raman imaging. *J Neurooncol.* 2015; 125(2):287-295.
185. Zhou Y, Liu CH, Sun Y, et al. Human brain cancer studied by resonance Raman spectroscopy. *Journal of Biomedical Optics.* 2012; 17(11):116021.
186. Aguiar RP, Silveira L, Jr., Falcao ET, Pacheco MT, Zangaro RA, Pasqualucci CA. Discriminating neoplastic and normal brain tissues in vitro through Raman spectroscopy: a

principal components analysis classification model. *Photomedicine and Laser Surgery.* 2013; 31(12):595-604.
187. Beljebbar A, Dukic S, Amharref N, Manfait M. Ex vivo and in vivo diagnosis of C6 glioblastoma development by Raman spectroscopy coupled to a microprobe. *Anal Bioanal Chem.* 2010; 398(1):477-487.
188. Kirsch M, Schackert G, Salzer R, Krafft C. Raman spectroscopic imaging for in vivo detection of cerebral brain metastases. *Analytical & Bioanalytical Chemistry.* 2010; 398(4):1707-1713.
189. Desroches J, Jermyn M, Mok K, et al. Characterization of a Raman spectroscopy probe system for intraoperative brain tissue classification. *Biomed Opt Express.* 2015; 6(7):2380-2397.
190. Jermyn M, Mok K, Mercier J, et al. Intraoperative brain cancer detection with Raman spectroscopy in humans. *Sci Transl Med.* 2015; 7(274):274ra219.
191. Desroches J, Lemoine E, Pinto M, et al. Development and first in-human use of a Raman spectroscopy guidance system integrated with a brain biopsy needle. *J Biophotonics.* 2019; 12(3):e201800396.
192. Fullwood LM, Griffiths D, Ashton K, et al. Effect of substrate choice and tissue type on tissue preparation for spectral histopathology by Raman microspectroscopy. *Analyst.* 2014; 139(2):446-454.
193. Nallala J, Lloyd GR, Stone N. Evaluation of different tissue de-paraffinization procedures for infrared spectral imaging. *Analyst.* 2015; 140(7):2369-2375.
194. Uckermann O, Yao W, Juratli TA, et al. IDH1 mutation in human glioma induces chemical alterations that are amenable to optical Raman spectroscopy. *J Neurooncol.* 2018.
195. Jermyn M, Desroches J, Mercier J, et al. Raman spectroscopy detects distant invasive brain cancer cells centimeters beyond MRI capability in humans. *Biomed Opt Express.* 2016; 7(12):5129-5137.
196. Uckermann O, Galli R, Tamosaityte S, et al. Label-free delineation of brain tumors by coherent anti-Stokes Raman scattering microscopy in an orthotopic mouse model and human glioblastoma. *PLoS ONE [Electronic Resource].* 2014; 9(9):e107115.
197. Evans CL, Xu X, Kesari S, Xie XS, Wong ST, Young GS. Chemically-selective imaging of brain structures with CARS microscopy. *Opt Express.* 2007; 15(19):12076-12087.
198. Camp CH, Jr., Lee YJ, Heddleston JM, et al. High-Speed Coherent Raman Fingerprint Imaging of Biological Tissues. *Nat Photonics.* 2014; 8:627-634.
199. Galli R, Uckermann O, Koch E, Schackert G, Kirsch M, Steiner G. Effects of tissue fixation on coherent anti-Stokes Raman scattering images of brain. *Journal of Biomedical Optics.* 2014; 19(7):071402.
200. Freudiger CW, Min W, Saar BG, et al. Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science.* 2008; 322(5909):1857-1861.
201. Freudiger CW, Pfannl R, Orringer DA, et al. Multicolored stain-free histopathology with coherent Raman imaging.[Erratum appears in Lab Invest. 2012 Nov;92(11):1661 Note: Ying, Wei [corrected to Wei, Ying]]. *Laboratory Investigation.* 2012; 92(10):1492-1502.
202. Ji M, Orringer DA, Freudiger CW, et al. Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy. *Sci Transl Med.* 2013; 5(201):201ra119.
203. Ji M, Lewis S, Camelo-Piragua S, et al. Detection of human brain tumor infiltration with quantitative stimulated Raman scattering microscopy. *Sci Transl Med.* 2015; 7(309):309ra163.
204. Orringer DA, Pandian B, Niknafs YS, et al. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nature Biomedical Engineering.* 2017; 1:0027.
205. Hollon TC, Lewis S, Pandian B, et al. Rapid Intraoperative Diagnosis of Pediatric Brain Tumors Using Stimulated Raman Histology. *Cancer Res.* 2018; 78(1):278-289.
206. Krafft C, Sobottka SB, Schackert G, Salzer R. Raman and infrared spectroscopic mapping of human primary intracranial tumors: a comparative study. *Journal of Raman Spectroscopy.* 2006; 37(1-3):367-375.
207. Kendall C, Isabelle M, Bazant-Hegemark F, et al. Vibrational spectroscopy: a clinical tool for cancer diagnostics. *Analyst.* 2009; 134(6):1029-1045.

208. Ali SM, Bonnier F, Lambkin H, et al. A comparison of Raman, FTIR and ATR-FTIR micro spectroscopy for imaging human skin tissue sections. *Analytical Methods.* 2013; 5(9):2281-2291.
209. Baker MJ, Byrne HJ, Chalmers J, et al. Clinical applications of infrared and Raman spectroscopy: state of play and future challenges. *Analyst.* 2018; 143(8):1735-1757.
210. Krafft C, Thummler K, Sobottka SB, Schackert G, Salzer R. Classification of malignant gliomas by infrared spectroscopy and linear discriminant analysis. *Biopolymers.* 2006; 82(4):301-305.
211. Krafft C, Shapoval L, Sobottka SB, Schackert G, Salzer R. Identification of Primary Tumors of Brain Metastases by Infrared Spectroscopic Imaging and Linear Discriminant Analysis. *Technology in Cancer Research & Treatment.* 2006; 5(3):291-298.
212. Uckermann O, Juratli TA, Galli R, et al. Optical analysis of glioma: Fourier-transform infrared spectroscopy reveals the IDH1 mutation status. *Clin Cancer Res.* 2017.
213. Mackanos MA, Contag CH. Fiber-optic probes enable cancer detection with FTIR spectroscopy. *Trends Biotechnol.* 2010; 28(6):317-323.
214. Ollesch J, Zaczek M, Heise HM, et al. Clinical application of infrared fibre-optic probes for the discrimination of colorectal cancer tissues and cancer grades. *Vibrational Spectroscopy.* 2017; 91:99-110.
215. Valdés PA, Roberts DW, Lu F-K, PhD, Golby A. Optical technologies for intraoperative neurosurgical guidance. *Neurosurgical Focus.* 2016; 40(3):E8.
216. Stummer W, Novotny A, Stepp H, Goetz C, Bise K, Reulen HJ. Fluorescence-guided resection of glioblastoma multiforme by using 5-aminolevulinic acid-induced porphyrins: a prospective study in 52 consecutive patients. *J Neurosurg.* 2000; 93(6):1003-1013.
217. Lau D, Hervey-Jumper SL, Chang S, et al. A prospective Phase II clinical trial of 5-aminolevulinic acid to assess the correlation of intraoperative fluorescence intensity and degree of histologic cellularity during resection of high-grade gliomas. *J Neurosurg.* 2016; 124(5):1300-1309.
218. Valdés PA, Kim A, Leblond F, et al. Combined fluorescence and reflectance spectroscopy for in vivo quantification of cancer biomarkers in low- and high-grade glioma surgery. *Journal of biomedical optics.* 2011; 16(11):116007-116007.
219. Valdes PA, Leblond F, Jacobs VL, Wilson BC, Paulsen KD, Roberts DW. Quantitative, spectrally-resolved intraoperative fluorescence imaging. *Sci Rep.* 2012; 2:798.
220. Valdés PA, Leblond F, Kim A, Wilson BC, Paulsen KD, Roberts DW. A spectrally constrained dual-band normalization technique for protoporphyrin IX quantification in fluorescence-guided surgery. *Optics Letters.* 2012; 37(11):1817-1819.
221. Valdes PA, Jacobs VL, Wilson BC, Leblond F, Roberts DW, Paulsen KD. System and methods for wide-field quantitative fluorescence imaging during neurosurgery. *Optics Letters.* 2013; 38(15):2786-2788.
222. Valdes PA, Jacobs V, Harris BT, et al. Quantitative fluorescence using 5-aminolevulinic acid-induced protoporphyrin IX biomarker as a surgical adjunct in low-grade glioma surgery. *J Neurosurg.* 2015; 123(3):771-780.
223. Plaha P, Livermore LJ, Voets N, Pereira E, Cudlip S. Minimally invasive endoscopic resection of intraparenchymal brain tumors. *World Neurosurg.* 2014.
224. Ma R, Coulter CA, Livermore LJ, Voets NL, Al Awar O, Plaha P. Endoscopy in Temporal Lobe Glioma and Metastasis Resection: Is There a Role? *World Neurosurg.* 2018; 117:e238-e251.
225. Choo J, Takeuchi K, Nagata Y, et al. Neuroendoscopic Cylinder Surgery and 5-Aminolevulinic Acid Photodynamic Diagnosis of Deep-Seated Intracranial Lesions. *World Neurosurg.* 2018.
226. Rapp M, Kamp M, Steiger HJ, Sabel M. Endoscopic-assisted visualization of 5-aminolevulinic acid-induced fluorescence in malignant glioma surgery: a technical note. *World Neurosurg.* 2014; 82(1-2):e277-279.
227. Kerr LT, Byrne HJ, Hennelly BM. Optimal choice of sample substrate and laser wavelength for Raman spectroscopic analysis of biological specimen. *Analytical Methods.* 2015; 7(12):5041-5052.
228. Byrne HJ, Baranska M, Puppels GJ, et al. Spectropathology for the next generation: quo vadis? *Analyst.* 2015; 140(7):2066-2073.

229. Jackson NS, Chaudhary N, lambkin H, Malkin A, Meade AD. Optimisation of Deparaffinizing and Baking Adherence Techniques for Breast Tissue on Calcium Fluoride Slides for FTIR and Raman Imaging. *Faraday Discussion, Royal Society of Chemistry*. Cambridge2016.
230. Faoláin EÓ, Hunter MB, Byrne JM, et al. Raman Spectroscopic Evaluation of Efficacy of Current Paraffin Wax Section Dewaxing Agents. *Journal of Histochemistry & Cytochemistry*. 2005; 53(1):121-129.
231. Huang N, Short M, Zhao J, et al. Full range characterization of the Raman spectra of organs in a murine model. *Optics Express.* 2011; 19(23):22892-22909.
232. Movasaghi Z, Rehman S, Rehman IU. Raman Spectroscopy of Biological Tissues. *Applied Spectroscopy Reviews.* 2007; 42(5):493-541.
233. De Gelder J, De Gussem K, Vandenabeele P, Moens L. Reference database of Raman spectra of biological molecules. *Journal of Raman Spectroscopy*. 2007; 38(9):1133-1147.
234. Li R, Verreault D, Payne A, et al. Effects of laser excitation wavelength and optical mode on Raman spectra of human fresh colon, pancreas, and prostate tissues. *Journal of Raman Spectroscopy*. 2014; 45(9):773-780.
235. R.L. M. *Raman Spectroscopy for Chemical Analysis*: Wiley and Sons, New York; 2005.
236. Lewis AT, Gaifulina R, Isabelle M, et al. Mirrored stainless steel substrate provides improved signal for Raman spectroscopy of tissue and cells. *J Raman Spectrosc.* 2017; 48(1):119-125.
237. Lieber CA, Mahadevan-Jansen A. Automated Method for Subtraction of Fluorescence from Biological Raman Spectra. *Applied Spectroscopy*. 2003; 57(11):1363-1367.
238. Hanlon EB, Manoharan R, Koo TW, et al. Prospects for in vivo Raman spectroscopy. *Phys Med Biol*. 2000; 45(2):R1-59.
239. Bonnier F, Mehmood A, Knief P, et al. In vitro analysis of immersed human tissues by Raman microspectroscopy. *Journal of Raman Spectroscopy*. 2011; 42(5):888-896.
240. Bonnier F, Ali SM, Knief P, et al. Analysis of human skin tissue by Raman microspectroscopy: Dealing with the background. *Vibrational Spectroscopy*. 2012; 61:124-132.
241. Xie R, Chung J-Y, Ylaya K, et al. Factors Influencing the Degradation of Archival Formalin-Fixed Paraffin-Embedded Tissue Sections. *Journal of Histochemistry & Cytochemistry*. 2011; 59(4):356-365.
242. Levine J, Tavella P, Santarelli G. Introduction to the Special Issue on Celebrating the 50th Anniversary of the Allan Variance. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2016; 63(4):511-512.
243. Beier BD, Berger AJ. Method for automated background subtraction from Raman spectra containing known contaminants. *Analyst*. 2009; 134(6):1198-1202.
244. Zhao J, Lui H, McLean DI, Zeng H. Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy. *Appl Spectrosc.* 2007; 61(11):1225-1232.
245. Redmond R. Handbook of Biomedical Fluorescence" Introduction to Fluorescence and Photophysics: Marcel Dekker, NewYork; 2003.
246. Grosserueschkamp F, Bracht T, Diehl HC, et al. Spatial and molecular resolution of diffuse malignant mesothelioma heterogeneity by integrating label-free FTIR imaging, laser capture microdissection and proteomics. *Sci Rep*. 2017; 7:44829.
247. Tonjes M, Barbus S, Park YJ, et al. BCAT1 promotes cell proliferation through amino acid catabolism in gliomas carrying wild-type IDH1. *Nat Med.* 2013; 19(7):901-908.
248. Afseth NK, Kohler A. Extended multiplicative signal correction in vibrational spectroscopy, a tutorial. *Chemometrics and Intelligent Laboratory Systems.* 2012; 117:92-99.
249. Lever J, Krzywinski M, Altman N. Principal component analysis. *Nature Methods.* 2017; 14:641.
250. Téllez Soto CA, Medeiros-Neto LP, dos Santos L, et al. Infrared and confocal Raman spectroscopy to differentiate changes in the protein secondary structure in normal and abnormal thyroid tissues. *Journal of Raman Spectroscopy*. 2018; 49(7):1165-1173.
251. Emir UE, Larkin SJ, de Pennington N, et al. Noninvasive Quantification of 2-Hydroxyglutarate in Human Gliomas with IDH1 and IDH2 Mutations. *Cancer Res*. 2016; 76(1):43-49.
252. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ. 2-Hydroxyglutarate MR spectroscopy for prediction of isocitrate dehydrogenase mutant glioma: a systemic review and meta-analysis using individual patient data. *Neuro Oncol.* 2018; 20(12):1573-1583.

253. Shankar GM, Francis JM, Rinne ML, et al. Rapid Intraoperative Molecular Characterization of Glioma. *JAMA Oncol.* 2015; 1(5):662-667.
254. Ohka F, Yamamichi A, Kurimoto M, et al. A novel all-in-one intraoperative genotyping system for IDH1-mutant glioma. *Brain Tumor Pathol.* 2017; 34(2):91-97.
255. Esmaeili M, Hamans BC, Navis AC, et al. IDH1 R132H mutation generates a distinct phospholipid metabolite profile in glioma. *Cancer Res.* 2014; 74(17):4898-4907.
256. Maus A, Peters GJ. Glutamate and alpha-ketoglutarate: key players in glioma metabolism. *Amino Acids.* 2017; 49(1):21-32.
257. Strickland M, Stoll EA. Metabolic Reprogramming in Glioma. *Front Cell Dev Biol.* 2017; 5:43.
258. Sasaki M, Knobbe CB, Itsumi M, et al. D-2-hydroxyglutarate produced by mutant IDH1 perturbs collagen maturation and basement membrane function. *Genes Dev.* 2012; 26(18):2038-2049.
259. Turcan S, Rohle D, Goenka A, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature.* 2012; 483(7390):479-483.
260. Branzoli F, Pontoizeau C, Tchara L, et al. Cystathionine as a marker for 1p/19q codeleted gliomas by in vivo magnetic resonance spectroscopy. *Neuro Oncol.* 2019.
261. Fack F, Tardito S, Hochart G, et al. Altered metabolic landscape in IDH-mutant gliomas affects phospholipid, energy, and oxidative stress pathways. *EMBO Mol Med.* 2017; 9(12):1681-1695.
262. Clemens G, Hands JR, Dorling KM, Baker MJ. Vibrational spectroscopic methods for cytology and cellular research. *Analyst.* 2014; 139(18):4411-4444.
263. Zhang Q, Zhang P, Gou H, et al. Towards high-throughput microfluidic Raman-activated cell sorting. *Analyst.* 2015; 140(18):6163-6174.
264. Song Y, Yin H, Huang WE. Raman activated cell sorting. *Curr Opin Chem Biol.* 2016; 33:1-8.
265. Dang L, White DW, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature.* 2009; 462(7274):739-744.
266. Walsh S, Diamond D. Non-linear curve fitting using microsoft excel solver. *Talanta.* 1995; 42(4):561-572.
267. Krammer B, Plaetzer K. ALA and its clinical impact, from bench to bedside. *Photochem Photobiol Sci.* 2008; 7(3):283-289.
268. EnzoLifeSciences. Protoporphyrin IX (free acid) Product data sheet. http://www.enzolifesciences.com/fileadmin/reports/Datasheet-ALX-430-041.pdf.
269. Krieg M, Whitten DG. Self-sensitized photooxidation of protoporphyrin IX and related free-base porphyrins in natural and model membrane systems. Evidence for novel photooxidation pathways involving amino acids. *Journal of the American Chemical Society.* 1984; 106(8):2477-2479.
270. König K, Wabnitz H, Dietel W. Variation in the fluorescence decay properties of haematoporphyrin derivative during its conversion to photoproducts. *Journal of Photochemistry and Photobiology B: Biology.* 1990; 8(1):103-111.
271. Bezdetnaya L, Zeghari N, Belitchenko I, et al. Spectroscopic and Biological Testing of Photobleaching of Porphyrins in Solutions*. *Photochemistry and Photobiology.* 1996; 64(2):382-386.
272. Moan J, Sommer S. Oxygen Dependence of the Photosensitizing Effect of Hematoporphyrin Derivative in NHIK 3025 Cells. *Cancer Research.* 1985; 45(4):1608-1610.
273. Mang TS, Dougherty TJ, Potter WR, Boyle DG, Somer S, Moan J. PHOTOBLEACHING OF PORPHYRINS USED IN PHOTODYNAMIC THERAPY AND IMPLICATIONS FOR THERAPY. *Photochemistry and Photobiology.* 1987; 45(4):501-506.
274. Robinson DJ, Bruijn HS, Veen N, Stringer MR, Brown SB, Star WM. Protoporphyrin IX Fluorescence Photobleaching during ALA-Mediated Photodynamic Therapy of UVB-Induced Tumors in Hairless Mouse Skin. *Photochemistry and Photobiology.* 1999; 69(1):61-70.
275. Bagdonas S, Ma L-W, Iani V, Rotomskis R, Juzenas P, Moan J. Phototransformations of 5-Aminolevulinic Acid–induced Protoporphyrin IX in vitro: A Spectroscopic Study¶. *Photochemistry and Photobiology.* 2000; 72(2):186-192.
276. Peng Q, Berg K, Moan J, Kongshaug M, Nesland JM. 5-Aminolevulinic acid-based photodynamic therapy: principles and experimental research. *Photochem Photobiol.* 1997; 65(2):235-251.

277. Kennedy JC, Pottier RH. New trends in photobiology: Endogenous protoporphyrin IX, a clinically useful photosensitizer for photodynamic therapy. *Journal of Photochemistry and Photobiology B: Biology.* 1992; 14(4):275-292.
278. Monteiro HP, Abdalla DSP, Faljoni-Alàrio A, Bechara EJH. Generation of active oxygen species during coupled autoxidation of oxyhemoglobin and δ-aminolevulinic acid. *Biochimica et Biophysica Acta (BBA) - General Subjects.* 1986; 881(1):100-106.
279. Monteiro HP, Abdalla DS, Augusto O, Bechara EJ. Free radical generation during delta-aminolevulinic acid autoxidation: induction by hemoglobin and connections with porphyrinpathies. *Arch Biochem Biophys.* 1989; 271(1):206-216.
280. Cox GS, Whitten DG. Mechanisms for the photooxidation of protoporphyrin IX in solution. *Journal of the American Chemical Society.* 1982; 104(2):516-521.
281. Ericson MB, Grapengiesser S, Gudmundson F, et al. A spectroscopic study of the photobleaching of protoporphyrin IX in solution. *Lasers in Medical Science.* 2003; 18(1):56-62.
282. *Dye Lasers*: Springer, Berlin, Heidelberg; 1990.
283. *Dye Laser Principles*: Elsevier, New York; 1990.
284. Hammond PR. Laser dye DCM, its spectral properties, synthesis and comparison with other dyes in the red. *Optics Communications.* 1979; 29(3):331-333.
285. Hammond PR. Self-absorption of molecular fluorescence, the design of equipment for measurement of fluorescence decay, and the decay times of some laser dyesa). *The Journal of Chemical Physics.* 1979; 70(8):3884-3894.
286. Marason EG. Laser dye DCM: CW, synchronously pumped, cavity pumped and single-frequency performance. *Optics Communications.* 1981; 37(1):56-58.
287. Drake J, Lesiecki M, Camaioni D. Photophysics and cis-trans isomerization of DCM. *Chem. Phys. Lett.* 1985; 113:530-534.
288. Gouterman M, Khalil G-E. Porphyrin free base phosphorescence. *Journal of Molecular Spectroscopy.* 1974; 53(1):88-100.
289. Kim A, Khurana M, Moriyama Y, Wilson BC. Quantification of in vivo fluorescence decoupled from the effects of tissue optical properties using fiber-optic spectroscopy measurements. *J Biomed Opt.* 2010; 15(6):067006.
290. Lilge L, O'Carroll C, Wilson BC. A solubilization technique for photosensitizer quantification in ex vivo tissue samples. *Journal of Photochemistry and Photobiology B: Biology.* 1997; 39(3):229-235.
291. Valdes PA, Kim A, Brantsch M, et al. delta-aminolevulinic acid-induced protoporphyrin IX concentration correlates with histopathologic markers of malignancy in human gliomas: the need for quantitative fluorescence-guided resection to identify regions of increasing malignancy. *Neuro Oncol.* 2011; 13(8):846-856.
292. Valdés PA, Leblond F, Kim A, et al. Quantitative fluorescence in intracranial tumor: implications for ALA-induced PpIX as an intraoperative biomarker. *J Neurosurg.* 2011; 115(1):11-17.
293. Panigrahi M, Gupta B, Reddy R. Neuroendoscopy - Is it safe? *Asian journal of neurosurgery.* 2017; 12(1):17-21.
294. Cahill DP, Sloan AE, Nahed BV, et al. The role of neuropathology in the management of patients with diffuse low grade glioma: A systematic review and evidence-based clinical practice guideline. *J Neurooncol.* 2015; 125(3):531-549.
295. Altman D, Machin D, Bryant T, Gardner(Eds) M. *Statistics with confidence*. 2 ed: BMJ Books; 2000.
296. Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis to case-control studies. *Stat Med.* 2007; 26(10):2170-2183.
297. CLOPPER CJ, PEARSON ES. THE USE OF CONFIDENCE OR FIDUCIAL LIMITS ILLUSTRATED IN THE CASE OF THE BINOMIAL. *Biometrika.* 1934; 26(4):404-413.
298. Kim S, Lee W. Does McNemar's test compare the sensitivities and specificities of two diagnostic tests? *Stat Methods Med Res.* 2017; 26(1):142-154.
299. Petrie A, Sabin C. *Medical Statistics at a Glance*. Third ed: Wiley-Blackwell; 2009.
300. Louis D, Ohgak H, Wiestler O, et al. *WHO classification of tumours of the central nervous system.* Revised 4th ed2016.

301. Meyer T, Bergner N, Bielecki C, et al. Nonlinear microscopy, infrared, and Raman microspectroscopy for brain tumor analysis. *Journal of Biomedical Optics*. 2011; 16(2):021113.
302. Advanced Sterilization Products. https://www.emea.aspjj.com/sterrad-systems.