

Titre: A New Representation for Spectral Data Applied to Raman
Title: Spectroscopy of Brain Cancer

Auteur: Émile Lemoine
Author:

Date: 2020

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Lemoine, É. (2020). A New Representation for Spectral Data Applied to Raman
Citation: Spectroscopy of Brain Cancer [Master's thesis, Polytechnique Montréal].
PolyPublie. <https://publications.polymtl.ca/5217/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/5217/>
PolyPublie URL:

**Directeurs de
recherche:** Frédéric Leblond, & Kevin Petrecca
Advisors:

Programme: Génie biomédical
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**A NEW REPRESENTATION FOR SPECTRAL DATA APPLIED TO
RAMAN SPECTROSCOPY OF BRAIN CANCER**

ÉMILE LEMOINE

Institut de génie biomédical

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie biomédical

Mars 2020

© Émile Lemoine, 2020.

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

A NEW REPRESENTATION FOR SPECTRAL DATA APPLIED TO RAMAN SPECTROSCOPY OF BRAIN CANCER

présenté par **Émile LEMOINE**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Julien COHEN-ADAD, président

Frédéric LEBLOND, membre et directeur de recherche

Kevin PETRECCA, membre et codirecteur de recherche

Jean PROVOST, membre

ACKNOWLEDGEMENTS

I would first like to thank my advisor Frédéric for his infinite support throughout this project. Frédéric believed in me the first day we met and has been pushing me further out of my comfort zone ever since. He allowed me to be creative and pursue my own ideas and solutions, while always being there to steer me on track if I needed it. His constant availability allowed to comfort my strongest concerns and doubts. He infused in me a strong and deep-rooted passion for research.

I would also like to thank my co-supervisor, Dr Kevin Petrecca, for his mentorship and advice in all aspects of this work. His dedication and scientific rigor have contributed to make me a better scientist. Most of all, he inspired me to always seek to improve the care of my future patients.

Beside my advisors, I want to express my sincere gratitude towards all the students at the Laboratory of Radiological Optics, who welcomed me as one of their own. Their help and patience would save me hours of frustrations in understanding the most basic concepts in optics and data science. Especially, I want to thank Frédérick, an incredible mentor and friend. Our endless discussions played a pivotal role in most of what I accomplished during these two years.

I would also like to acknowledge the help of Damon, Katherine and all other co-authors with whom I have had the chance to collaborate on several projects.

Lastly, I would like to express my deepest gratitude to Justine and the rest of my family. This thesis would simply not have been possible without their understanding, support and love.

RÉSUMÉ

Par sa nature infiltrative et son confinement derrière la barrière hémato-encéphalique, le cancer primaire du cerveau est l'une des néoplasies les plus difficiles à diagnostiquer et traiter. Son traitement repose sur la résection chirurgicale maximale. La spectroscopie Raman, capable d'identifier en temps réel des régions cancéreuses qui apparaîtraient normales à l'œil nu, promet d'améliorer considérablement le guidage neurochirurgical et maximiser la résection de la masse tumorale. Cependant, le signal Raman est très complexe à interpréter : les systèmes Raman peuvent maintenant capter des signaux de grande qualité que les méthodes analytiques actuelles ne parviennent pas à interpréter de manière reproductible. Ceci constitue une barrière importante à l'acceptation de la spectroscopie Raman par les médecins et les chercheurs œuvrant sur le cancer du cerveau.

L'objectif de ce travail est de développer une méthode robuste d'ingénierie des variables (« *Feature engineering* ») qui permettrait d'identifier les processus moléculaires exploités par les systèmes Raman pour différencier les régions cancéreuses des régions saines lors de chirurgies cérébrales.

Tout d'abord, nous avons identifié les régions Raman ayant une haute spécificité à notre problématique clinique par une revue systématique de la littérature. Un algorithme d'ajustement de courbe a été développé afin d'extraire la forme des pics Raman dans les régions sélectionnées. Puis, nous avons élaboré un modèle mathématique qui tient compte de l'interactivité entre les molécules de l'échantillon interrogé, ainsi qu'entre le signal Raman et l'âge du patient opéré. Pour valider le modèle, nous avons comparé sa capacité à compresser le signal avec celle de l'analyse en composante principale (ACP), le standard en spectroscopie Raman. Finalement, nous avons appliqué la méthode d'ingénierie des variables à des spectres Raman acquis en salle d'opération afin d'identifier quels processus moléculaires indiquaient la présence de cancer.

Notre méthode a démontré une meilleure rétention d'information que l'ACP. En l'appliquant aux spectres Raman *in vivo*, les zones denses en cellules malignes démontrent une expression augmentée d'acides nucléiques ainsi que de certaines protéines, notamment le collagène, le tryptophan et la phénylalanine. De plus, l'âge des patients semble affecter l'impact qu'ont certaines protéines, lipides et acides nucléiques sur le spectre Raman. Nos travaux révèlent l'importance d'une modélisation statistique appropriée pour l'implémentation clinique de systèmes Raman chirurgicaux.

ABSTRACT

Because of its infiltrative nature and concealment behind the blood-brain barrier, primary brain cancer remains one of the most challenging oncological condition to diagnose and treat. The mainstay of treatment is maximal surgical resection. Raman spectroscopy has shown great promise to guide surgeons intraoperatively by identifying, in real-time, dense cancer regions that appear normal to the naked eye. The Raman signal of living tissue is, however, very challenging to interpret, and while most advances in Raman systems targeted the hardware, appropriate statistical modeling techniques are lacking. As a result, there is conflicting evidence as to which molecular processes are captured by Raman probes. This limitation hinders clinical translation and usage of the technology by the cancer-research community.

This work focuses on the analytical aspect of Raman-based surgical systems. Its objective is to develop a robust data processing pipeline to confidently identify which molecular phenomena allow Raman systems to differentiate healthy brain and cancer during neurosurgeries.

We first selected high-yield Raman regions based on previous literature on the subject, resulting in a list of reproducible Raman bands with high likelihood of brain-specific Raman signal. We then developed a peak-fitting algorithm to extract the shape (height and width) of the Raman signal at those specific bands. We described a mathematical model that accounted for all possible interactions between the selected Raman peaks, and the interaction between the peaks' shape and the patient's age. To validate the model, we compared its capacity to compress the signal while maintaining high information content against a Principal Component Analysis (PCA) of the Raman spectra, the fields' standard. As a final step, we applied the feature engineering model to a dataset of intraoperative human Raman spectra to identify which molecular processes were indicative of brain cancer.

Our method showed better information retention than PCA. Our analysis of *in vivo* Raman measurement showed that areas with high-density of malignant cells had increased expression of nucleic acids and protein compounds, notably collagen, tryptophan and phenylalanine. Patient age seemed to affect the impact of nucleic acids, proteins and lipids on the Raman spectra. Our work demonstrates the importance of appropriate statistical modeling in the implementation of Raman-based surgical devices.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	III
RÉSUMÉ.....	IV
ABSTRACT	V
TABLE OF CONTENTS	VI
LISTE OF TABLES.....	VIII
LIST OF FIGURES.....	IX
LIST OF SYMBOLS AND ABBREVIATIONS.....	XI
CHAPTER 1 INTRODUCTION.....	1
CHAPTER 2 OVERALL APPROACH AND ORGANIZATION OF THE MEMOIRE	4
2.1 Organization of the memoire.....	5
CHAPTER 3 LITERATURE REVIEW.....	6
3.1 Article 1: The Rise of Raman spectroscopy in neurosurgery.....	6
3.1.1 The Rise of Raman spectroscopy in neurosurgery.....	6
3.2 Statistical Learning.....	49
3.2.1 Statistical models.....	49
3.2.2 Overfitting, underfitting and regularization	54
3.2.3 Frequentist and Bayesian inference.....	56
CHAPTER 4 ARTICLE 2: FEATURE ENGINEERING APPLIED TO INTRAOPERATIVE IN VIVO RAMAN SPECTROSCOPY SHEDS LIGHT ON MOLECULAR PROCESSES IN BRAIN CANCER: A RETROSPECTIVE STUDY OF 65 PATIENTS	65
4.1 Feature Engineering Applied to Intraoperative In Vivo Raman Spectroscopy Sheds Light on Molecular Processes in Brain Cancer: A Retrospective Study of 65 Patients	65
4.1.1 Abstract	66
4.1.2 Introduction	66

4.1.3	Methods.....	68
4.1.4	Results.....	77
4.1.5	Discussion.....	86
4.1.6	Conclusion.....	91
4.1.7	Conflicts of Interest.....	92
4.1.8	Acknowledgements.....	92
4.1.9	Tables.....	93
4.1.10	Supplementary information.....	94
CHAPTER 5	GENERAL DISCUSSION.....	102
5.1	Limitations.....	103
5.2	Importance.....	104
CHAPTER 6	FUTURE IMPLICATIONS AND CONCLUSION.....	106
BIBLIOGRAPHIE	108

LISTE OF TABLES

Table 3.1: Optical characteristics of neurosurgical point-probes.	44
Table 3.2: System characteristics of neurosurgical point-probes.	45
Table 3.3: Optical Characteristics of state-of-the-art imaging systems for neurosurgery.....	46
Table 3.4: Physical Characteristics of state-of-the-art imaging systems for neurosurgery.	47
Table 3.5: Feature engineering and supervised machine learning algorithms for classification of Raman spectroscopy in brain tissue	48
Table 4.1 : Description of selected, high-quality tissue samples used for the Raman acquisitions	93
Table 4.2 : Raman bands selected as target for the band fitting procedure, with molecular assignment based on literature.	93
Supplementary Table 4.1 : Raman peaks reported in literature on brain tissue from 1999 to 2019.	97

LIST OF FIGURES

Figure 3.1: Stages of translational systems for neurosurgery.	15
Figure 3.2: SRS contrast in brain tissue.	19
Figure 3.3: Translational Raman imaging systems.	23
Figure 3.4: Spectral processing of raw Raman signal in biological tissue.....	27
Figure 3.5: Machine learning for Raman spectroscopy.	29
Figure 3.6: Spectral unmixing and clustering-based approaches to Raman imaging, where every pixel consists in an entire processed Raman spectrum.	33
Figure 3.7: Chemometrics analysis of Raman data.....	37
Figure 3.8: Linear model for a dataset with a single variable x and an outcome y . The model has the form $y = \alpha + \beta x$	51
Figure 3.9: Modeling binary outcome with one predictor variable.....	52
Figure 3.10: Simulated example to illustrate the effect of an interaction term in a linear model.	53
Figure 3.11: Depiction of the posterior distribution as the result of the prior distribution being updated by the likelihood, or collected data.....	58
Figure 3.12: Bayesian Hyper-Lasso priors.....	59
Figure 3.13: Posterior predictive distribution for a single observation x_i	63
Figure 4.1 : Details of the feature engineering method.....	73
Figure 4.2 : Depiction of the workflow used for the acquisition of the Raman spectroscopy data.	78
Figure 4.3 : Prevalent peaks in literature vs. prevalent peaks in the dataset.	79
Figure 4.4 : Extraction of an interpretable representation for Raman data.	81
Figure 4.5 : Comparison of reconstruction error and cancer predictive performances between the Raman representation resulting from band fitting and dimensionality reduction following Principal Component Analysis (PCA).	82

Figure 4.6 : Raman spectral features from the new representation selected as probably (>75%) informative of the presence of dense cancer in a sample by the multivariate analysis when controlling for peak width.	84
Figure 4.7 : Interactive effect of age on selected Raman bands.	85
Supplementary Figure 4.1 : Averaged Raman spectra with standard deviation acquired in human brain samples during neurosurgical glioma resection, stratified by acquisition time periods.	98
Supplementary Figure 4.2 : Averaged Raman acquisitions before (“Raw spectra”) and after data processing.	99
Supplementary Figure 4.3 : All individual Raman acquisitions ($n = 547$) stratified by data quality.	100
Supplementary Figure 4.4 : All Raman bands mentioned in the literature on Raman spectroscopy of brain tissue from 1999–2019.	101

LIST OF SYMBOLS AND ABBREVIATIONS

5-ALA	5-aminolaevulinic acid
AF	Autofluorescence
ANN	Artificial neural network
CARS	Coherent anti-Stokes Raman scattering
CCD	Charge-coupled device
CR	Coherent Raman
CT	Computed tomography
CW	Continuous wave
DBS	Deep brain stimulation
DRS	Diffuse reflectance spectroscopy
FFPP	Formalin fixed paraffin processed
FOV	Field-of-view
H&E	Hematoxylin and eosin
HWN	High wavenumber
LDA	Linear discriminant analysis
LDF	Laser Doppler Flow
LPD	Log-predictive density
MCMC	Markov Chain Monte-Carlo
ML	Machine learning
MLE	Maximum likelihood estimation
MPE	Maximal permissible exposure
MRI	Magnetic Resonance Imaging
MSE	Mean-squared-error

NAD	Nicotinamide adenine dinucleotide
NIR	Near-infrared
NRMSE	Normalised Root-mean-squared-error
OCT	Optical Coherence Tomography
OD	Outer diameter
PC(A)	Principal component (analysis)
PD	Parkinson's Disease
PET	Positron Emission Tomography
RS	Raman spectroscopy
SHG	Second harmonic generation
SNR	Signal-to-noise ratio
SNV	Standard normal variate
SRG(L)	Stimulated Raman gain (loss)
SRS	Stimulated Raman scattering
SVM	Support vector machine
TPEF	Two-photon excited fluorescence
VCA	Vertex component analysis

CHAPTER 1 INTRODUCTION

Cancer is projected to become the leading cause of death of the 21st century[1]. The rise of the cancer burden is counterbalanced by massive efforts and investments in discovering novel prevention, diagnosis and therapeutic strategies. Since the description of the *Hallmarks of Cancer* by Hanahan and Weinberg in 2000[2], the recognition of immunology, metabolomics and genetics as major determinants of cancer emergence and survival revolutionized our approach to this complex and heterogenous disease[3].

Among all cancer types, primary brain cancer remains one of the most challenging to diagnose and treat. A diagnosis of glioma—the most common form of malignant brain tumor—carries a median survival of five years. In the case of glioblastoma, the most aggressive and prevalent subtype of glioma, most patients die within the following 12 months[4], [5]. Their enclosed location and isolation from systemic circulation by the Blood-Brain barrier hinders treatment, the mainstay of which is the maximal resection of the tumor. However, new insights in brain cancer biology and microenvironment are leading to novel therapeutic options.

In the past, glioma was classified according to specific histological features such as necrosis and micro-angiogenesis. In 2016, a revision of World Health Organization brain tumor classification put increased emphasis on genetic and biologic markers, significantly impacting patient care[6]. The metabolism of tumor cells and their surrounding tissue is now recognized as having both therapeutic and prognostic implications. These metabolic markers can be detected either in a neuropathology laboratory, with extensive processing of the resected tissue, or with advanced imaging modalities, non-invasively and directly in the patients' brain. Yet, these highly specialized technologies (*e.g.* Magnetic Resonance Imaging/Spectroscopy [MRI/MRS], Positron Emission Tomography [PET]) require facilities and personnel, time, a lot of money, and, in some cases, the injection of exogeneous radiotracers[7]. Moreover, by imaging the brain inside the skull, they make concurrent analysis with spatially correlated tissue samples challenging[8]. *In vivo* glioma metabolomics is a promising field for diagnosis and therapy, but accessible modalities are lacking.

A brain cancer surgery offers a rare opportunity to come face-to-face with the living cancer cells. Its objective is to resect a maximal amount of cancer cells, while leaving the healthy tissues intact to minimize neurological deficits[5]. In recent years, great engineering efforts have been put into developing tools to guide resection of brain tumors and help surgeons detect the intractable cancer

margin[9]. Of those, Raman spectroscopy has emerged as one of the most promising techniques. This label-free, highly sensitive optical technology is used since the 1930s by chemists and biologists to study the vibrational profile of materials, cells and tissues. It is now being integrated into surgical tools to assist navigation inside the brain.

The Raman effect, an inelastic scattering of photons in matter, is a rare phenomenon: its intensity is about 1/10 million that of the emitting source. Given an incident light with spatial frequency ν_0 , the inelastically scattered photons will reach the detector with frequency ν' . The photons' shift in frequency $\Delta\nu = \nu' - \nu_0$, or *Raman shift*, depends on the vibrational profile of the molecules encountered during their trajectory. The distribution of the intensity of the detected signal as a function of its Raman shift gives rise to a Raman spectrum. The complexity of the Raman scattering depends on that of the interrogated sample. Biological tissues are rich and dynamic systems, composed of thousands of interacting molecules. Their structure and behavior are dictated by various co-founding factors. The resulting Raman signal is therefore extremely challenging to interpret[10].

Fortunately, following the recent explosion in computational capacities, our capacity to model complex data has greatly increased. Graph network analysis, reinforcement learning, hierarchical modeling, deep learning and Bayesian statistics have become the mainstay for many data-driven subfields in biology: neuroscience[11]–[14], genetics[15]–[18], oncology[19], and metabolomics[20], [21].

Nonetheless, most of the efforts in translating Raman spectroscopy to the clinic have focused on the hardware[22]. In the past 20 years, many research groups have shown that Raman systems can acquire reliable signal with minimal noise and artefacts, whether in the blood vessels[23], [24], gut[25]–[28], lungs[29], [30], breast[31]–[34], skin[35]–[37], genitourinary system[38]–[42], or brain[43]–[46]. But statistical analysis has stalled; it is poorly fit to high-dimensional, hierarchical, dynamic and interaction-rich data-generating processes[47]–[49]. As a result, there is conflicting evidence on which molecular processes and biomarkers really are captured by the technique—we can acquire a Raman spectrum in the brain but cannot reliably tell what it represents[22]. This poses important limitations on the use of Raman spectroscopy. For scientists, understanding which molecular phenomenon they are probing is important to generate meaningful hypothesis and widen the scope of the technology. For physicians, knowing how the probe “thinks”—how it arrives at a

specific diagnosis for an interrogated tissue—is crucial before accepting it as part of their clinical workflow. Unsurprisingly, statistical methods and thorough understanding of their underlying mechanisms are recognized as major factors for the successful clinical translation of a technology[50].

There is a gap to bridge between Raman spectroscopy and the novel methodologies to analyze high-dimensional data. This project is driven by the following question: can advanced analytical methods for high-dimensional data be leveraged to robustly identify which molecular processes allow Raman systems to differentiate cancer from normal? Accordingly, we developed a complete data analysis pipeline, starting from the processed Raman spectra, that is specifically designed for the clinical problem, informed by prior domain specific knowledge, and where robustness and uncertainty are explicitly modeled to predict future replicability. This algorithm is then applied to a large dataset of intraoperative Raman acquisitions, allowing a rich insight into the possible hidden driving forces of the Raman signal in cancer and normal tissue. This work demonstrates how, with proper statistical models, Raman spectroscopy can overcome its limitations and offer a window on living cancer cells[51].

CHAPTER 2 **OVERALL APPROACH AND ORGANIZATION OF THE MEMOIRE**

In the following chapters, we attempt to demonstrate the promising nature of Raman spectroscopy for the analysis of brain tissue, in both the research and clinical settings (neuropathology or neurosurgery). The technology, however, suffers from an important pitfall that limit its clinical translation and acceptability by scientists outside the field of bio-spectroscopy: Raman data from biological samples are high-dimensional and convoluted. Overcoming this challenge would greatly enhance the scope of this technology.

This memoire aims to answer the following question: what are we really capturing with Raman-based systems in living brain tissue? More specifically, in the setting of neurosurgical intra-operative tumor detection:

- 1) Which features of the Raman signal are useful to differentiate cancer from its surrounding tissue?
- 2) Do second-order interactions between Raman markers, and interactions between Raman signal and patient specific variables have a role in the tissue recognition capacities?

These questions have not been be answered using traditional data modeling approaches but could be resolved with newer statistical techniques. The objectives of this project were therefore:

- 1) To develop a robust method that transforms the Raman signal into a representation that is easily interpretable and amenable to functionally significant statistical analysis while limiting the loss of information compared to currently used methods;
- 2) To integrate interactions between Raman features in the analysis and include interactions between the Raman signal and a patient-specific factor. Because of its well-demonstrated effect on the molecular composition of the brain and brain cancer, age was selected as the patient-specific factor;
- 3) To apply this method to a dataset of high-quality, intraoperative Raman acquisitions of living brain samples and infer which molecular processes differentiate cancerous tissue from its surroundings.

2.1 Organization of the memoire

The memoire is divided in two main parts: in the first section (Chapter 3), the current literature is critically reviewed to demonstrate the capacities and promises held by Raman-based systems in neurosurgery and neuropathology, while exposing how the current analytical methods limit its potential applicability. Chapter 3.2 aimed to lay down the theoretical ground to important concept in statistical learning, particularly with regards to inference and Bayesian estimation, with emphasis on how it could potentially overcome challenges inherent to Raman data.

The second section of the memoire presents the methodology and results of the experiments (Chapter 4), followed by a brief return on the objectives and a discussion regarding limitations, hypotheses that this work has generated and future steps in the applications of feature engineering for Raman spectroscopy in neurosurgery (Chapters 5 and 6).

CHAPTER 3 LITERATURE REVIEW

The literature review is divided in two parts. The first part is a reproduction of a manuscript submitted for publication that presents Raman spectroscopy and its applications to neurosurgery. It highlights the important advances that rendered Raman amenable to the operating room, with an emphasis on the remaining obstacles to its clinical translation. It demonstrates how current data analysis techniques are not adapted to the complexity of *in vivo* Raman spectroscopy, resulting in a poor comprehension of the specific molecular processes targeted by the technique. The second part of the review focuses on statistical theory and introduces principles of Bayesian inference, and how these can help better model Raman data in the clinical setting.

3.1 Article 1: The Rise of Raman spectroscopy in neurosurgery

The following review article titled “The Rise of Raman spectroscopy in neurosurgery” was accepted for publication in the *Journal of Biomedical Optics* in April 2020. The manuscript was co-first authored by graduate student Damon DePaoli and Émile Lemoine, with contribution from postdoctoral fellow Katherine Ember, neurosurgeons Léo Cantin, Michel Prud’homme and Kevin Petrecca, and professors Martin Parent, Daniel Côté and Frédéric Leblond. The contribution of the memoire’s author includes the research and redaction for the sections that are most relevant to this memoire: Raman theory, neurosurgical approaches, advances in data analysis, and promising outlook. His contribution also includes participation in the overall design of the review and researching and reviewing the other sections.

This publication was reproduced with the permission of the Journal of Biomedical Optics in the following pages. References are included at the end of the memoire.

3.1.1 The Rise of Raman spectroscopy in neurosurgery

Journal: *Journal of Biomedical Optics*

Damon DePaoli^{1,2*} & Émile Lemoine^{3,4*}, Katherine Ember^{3,4}, Martin Parent¹, Michel Prud’homme⁵, Léo Cantin⁵, Kevin Petrecca⁶, Frédéric Leblond^{3,4}, Daniel C. Côté^{1,2}

¹CERVO Brain Research Center, Université Laval, Québec, Canada

²Centre d’optique, photonique et Lasers (COPL), Université Laval, Québec, Canada

³Department of Engineering Physics, Polytechnique Montréal, Montréal, Canada

⁴Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Canada

⁵Hôpital de l'Enfant-Jésus, Neurosurgery, Québec, Canada

⁶Department of Neurology and Neurosurgery, Montreal Neurological Institute-Hospital, McGill University

*These authors contributed equally to this work

3.1.1.1 Abstract

Significance: While the clinical potential for Raman spectroscopy (RS) has been anticipated for decades, it has only recently been used in neurosurgery. Still, few devices have succeeded to make their way into the operative room. With recent technological advancements, however, vibrational sensing is poised to be a revolutionary tool for neurosurgeons. **Aim:** This Perspective will give an overview of neurosurgical workflows, summarize key translational milestones of RS in clinical use, and provide the optics and data science background required to implement such devices. **Approach:** We performed an extensive review of the Raman literature, with a specific emphasis on research that aims to build Raman systems suited for the clinical setting. **Results:** The main translatable interest in Raman sensing rests in its capacity to yield label-free molecular information from tissue intraoperatively. Systems that have proven usable in the clinical setting are ergonomic, have a short integration time, and can acquire high-quality signal even in suboptimal conditions. Moreover, because of the complex microenvironment of brain tissue, data analysis is now recognized as critical to the performances of Raman-based surgical probes. **Conclusions:** The next generation of Raman-based devices are making their way into operating rooms, and their clinical translation requires close collaboration between physicians, engineers and data scientists.

3.1.1.2 Introduction

Neurosurgery can be used to treat a multitude of disorders ranging from brain tumors and cancers, to traumatic brain injury, epilepsy, and Parkinson's disease. 13.8 million neurosurgical procedures are carried out worldwide every year, and it is estimated that an additional 5 million neurosurgical conditions go untreated annually.[52] Neurosurgeons face many challenges that are unmet by

modern surgical techniques: incomplete tumor resection, inaccurate surgical guidance, expensive and inefficient intraoperative diagnostics, and a relatively high-risk of adverse events.

Optical technologies have gained considerable traction in neurosurgery over the last few decades. The use of 5-aminolevulinic acid induced cancer fluorescence in glioblastoma surgery for margin detection has become common.[53], [54] More recently, a variety of other techniques are being investigated for tumor margin detection including quantitative exogenous fluorescence,[55], [56] endogenous fluorescence lifetime imaging[57]–[59], optical coherence tomography (OCT),[60] hyperspectral imaging[61], and Raman spectroscopy (RS).[43] Optical techniques are also showing promise when applied to other neurosurgical procedures. For example, deep brain stimulation (DBS) for Parkinson’s disease (PD) could be optically guided using laser doppler flow (LDF) measurements[62], diffuse reflectance spectroscopy (DRS)[63]–[65], and coherent Raman (CR) spectroscopy.[66] During epilepsy surgery, hyperspectral imaging is being investigated to help guide resection.[67] Finally, in closed biopsies, OCT has been used to image blood vessels to minimize hemorrhage rates[68] and RS has shown promise in effective tumor targeting.[45], [69]

Here we focus on the potential of RS to improve several neurosurgical workflows. RS is an advantageous modality for biomedical applications because it can provide label-free, molecular-specific, information from tissue within safe limits of optical power. By analyzing this information with effective data science models, RS can be used to provide real-time discriminatory feedback and guidance to neurosurgeons. RS-based tools could be used to discriminate tumor and non-tumor tissue for cancer resection, to detect blood vessels for safe biopsy acquisition, and to detect novel biomarkers for disease diagnosis. However, there remain a number of challenges when applying RS clinically. The Raman scattering effect is a weak phenomenon and so interference from other optical processes can affect the measurable signal. Moreover, analysis of large spectral datasets obtained from complex biological mixtures can be a daunting and time-consuming task. Herein, we provide an outline of RS utility during neurosurgical procedures, a summary of spontaneous and coherent RS techniques, and a thorough review of the leading-edge systems and data analysis techniques already being deployed in, or in development for, a neurosurgical setting. All of this is provided so that future endeavors in this arena can be undertaken with clear objectives.

3.1.1.3 Clinical challenges in neurosurgery

The primary objectives of neurosurgeries involving brain tumors are: 1) procuring quality biopsy tissue for accurate diagnosis and 2) achieving maximal cancer resection while minimizing injury to the normal brain.[5], [70] Extensive multimodal imaging (e.g. Magnetic resonance imaging (MRI), X-ray computed tomography (CT) scans, positron emission tomography (PET) scans) is used preoperatively to characterize the location of the tumor, its relationship within the brain, and the imaging features of the tumor. These images are also used by Neuronavigation tracking devices during surgery. However, due to poor resolution and sensitivity, none of the preoperative imaging technique can visualize the full extent of invasive brain cancer, limiting surgical planning. In addition, Neuronavigation is highly susceptible to the shifting of the brain once the dura has been opened, and tumor and cerebrospinal fluid are removed.[71], [72]

For certain brain lesions a closed biopsy for tissue diagnosis is most appropriate. Studies have shown that brain biopsies may be non-diagnostic in up to 10% of cases, and even if a diagnosis is achieved, it is found to be inexact in as many as 23% of cases.[73]

In functional neurosurgery, the aim is to relieve patients suffering from chronic neurological or neurodegenerative disorders such as Parkinson's disease (PD), chronic pain, epilepsy, or dystonia. Deep brain stimulation (DBS) consists of surgical implantation of electrodes deep inside the brain where they modulate specific brain nuclei to correct dysfunctional brain circuits. Whilst many conditions can benefit from DBS surgery, its most common application is for alleviating motor symptoms of PD. The main challenge of DBS is the accurate positioning of the electrode inside the target nuclei.[74] Again, electrode placement accuracy is dictated by the quality of preoperative imaging used and the precision of the Neuronavigation tracking devices.

RS has shown great promise in overcoming these challenges. In brain tumor surgery it provides real-time, molecularly specific, information to neurosurgeons, accurately predicting the nature of the probed tissue prior to its resection. For closed brain biopsies, it can increase diagnostic yield and minimize harm by targeting cancer tissue before the biopsy sample is harvested. And, in functional neurosurgery, RS can guide electrode placement to improve safety, accuracy and clinical outcomes.

3.1.1.4 Raman spectroscopy techniques: a short primer

Whilst it is not the objective of this review to provide theory on the Raman effect, it is important to have an idea of the signal generation with respect to the sample and the optical power used. Furthermore, the wavelength of the generated signal can be an important design factor given operating room lighting, microscope illumination, and background signals. Lastly, spectral content is a critical aspect of the diagnostic ability associated with a clinical system and so the techniques for spectral acquisition are also discussed. For more complete details of Raman theory, abundant resources exist.[75]–[78] The modalities that have been used in brain studies include Spontaneous Raman (SR) scattering, Coherent anti-Stokes Raman scattering (CARS), and Stimulated Raman scattering (SRS).

3.1.1.4.1 *Spontaneous Raman scattering*

Since its discovery by Sir C. V. Raman in 1928, Raman scattering of light has been widely adopted as a molecular probing tool in the fields of biology and chemistry. The main feature of RS is the concept of vibrational energy: molecules in a sample vibrate. At room temperature, most molecules are in their ground state, i.e. the lowest energy level. When excited by electromagnetic radiation, the molecules will either absorb or scatter the excitation photons. Most of the scattering is elastic (Rayleigh scattering) where the molecule is transiently raised to a virtual energy level by a photon of specific energy and almost immediately returns to the ground state by emitting a photon of the same energy. In approximately one out of a million of interaction events, however, the energy of the emitted photons will have changed relative to the incident light. This phenomenon is called the Raman effect, or inelastic scattering. If the inelastically scattered photons are of lower energy than the exciting light source, this is known as Stokes scattering. If the scattered photons have gained energy relative to the incident light, this is known as anti-Stokes scattering. The Raman spectrum is a mapping of the intensity of scattered light as a function of its shift in frequency, or Raman shift. It is in essence a vibrational profile of the molecules present in the interrogated sample subject to the partial volume effect.[79]

In SR, the detected spectra are a linear combination of signals from all the molecules in the illuminated sample. Furthermore, the Raman intensity increases linearly as a function of excitation power and exposure time, facilitating spectral analysis. SR suffers from two limitations in biological tissue: low Raman signal (in absolute terms and relative to background); and high

contamination from autofluorescence (AF) signals from tissue and instrument components (*i.e.* fiber optics). Most vibrational spectroscopy techniques aim to overcome these challenges.[80]

3.1.1.4.2 Coherent anti-Stokes Raman scattering

CARS is a two-photon microscopy technique that uses two excitation wavelengths—a “pump” and a “Stokes” beam. When the difference between the two excitation frequencies is equivalent to a target vibrational mode, resonance occurs, generating strong non-linear anti-Stokes signal. Although it is less practical as a means to acquire data at macroscopic scales, CARS has several advantages over SR imaging: 1) it is orders of magnitude more sensitive at a specific vibration, 2) it provides intrinsic optical sectioning due to non-linear signal generation, and 3) signal generation is blue-shifted, removing the need for single-photon AF removal.[81]

In 1999, Zumbusch from the Xie group revived CARS imaging for biological tissues.[82] Subsequently, groups have continued to apply CARS to a multitude of biological structures, where it provided contrast from vibrations in DNA, lipids, proteins and water.[83]–[86] In 2005, Wang *et al.* presented the first imaging of myelin in *ex vivo* guinea pig spinal cord.[87] Soon after, the landmark study by Evans *et al.* reported CARS imaging of *in vivo* tissue at video rate speeds,[88] and since then, many others have followed in refining and optimizing CARS for *in vivo* imaging of nervous tissue.[89]–[94]

The signal generation for CARS is proportional to the quadratic intensity of the pump field multiplied by the Stokes field. CARS also scales quadratically with the number of oscillators in the sampled volume, making it specifically useful for interrogating high-density substances. In brain tissue, myelin (wrapped around many axons) is the main contrast agent due to its abundance of CH₂ moieties.

While coherent Raman (CR) scattering is generated at the both the Stokes (coherent Stokes Raman scattering, CSRS) and anti-Stokes frequency (CARS), the anti-Stokes signal is more commonly detected as it is stronger and unaffected by AF.

The main source of background signal in CARS is known as the non-resonant background; it is independent of the Raman shift and the excitation wavelength. Due to this effect, CARS is limited in sensitivity to sensing only high concentrations of molecules, in contrast to stimulated Raman scattering which is unaffected by non-resonant background sources.

3.1.1.4.3 Stimulated Raman scattering (SRS)

Just under a decade after the Xie group introduced CARS to biological imaging in 2008, Freudiger *et al.* of the same group presented fast stimulated Raman spectroscopy (SRS) imaging of biological tissue *in vivo*.^[95] Whilst this was not the first use of SRS in microscopy, it was the first study to use optical powers that were safe for live animal imaging, enabled by phase-sensitive lock-in detection. The potential for SRS was rapidly demonstrated by several groups, displaying an ability to image DNA mitosis,^[96] protein dynamics,^[97] and even measure neurotransmitter concentrations.^[98] , Due to its increased sensitivity over CARS for molecules at low concentrations, SRS is ideal for imaging nuclear contrast in neuropathology.^[99], ^[100]

As for CARS, the signal generation for SRS is non-linear. However, the SRS signal is proportional to the product of the intensity from the pump field and the Stokes field. Furthermore, unlike CARS, SRS has a linear dependence with oscillators density, making the detected signal easier to correlate to molecular concentrations within the sample. The wavelength of signal generation for SRS can correspond to a stimulated Raman gain (SRG) or a stimulated Raman loss (SRL) depending on which frequency is the probe (i.e. which is modulated for detection). Laser noise, shot noise, and electronic noise are all sources of noise in SRS imaging. They are especially important to minimize as small changes in the excitation laser must be measured due to the modulated aspect of SRS. The main sources of background for SRS are Raman-independent pump-probe effects. These include transient absorption, cross-phase modulation, and photothermal effects none of which are observed in CARS as the signal is generated at a new wavelength.^[72] A theoretical review of the signal to noise ratio (SNR) for CR techniques is provided by Min *et al.*^[77]

3.1.1.4.4 Spectral imaging with coherent Raman techniques

Coherent Raman imaging has historically been used for single frequency imaging, creating contrast from the coherent vibration of only a few molecular bonds. However, this is limited in spectral information and therefore also in diagnostic capability. In an effort to achieve rapid, hyperspectral image acquisition, strategies to produce CR spectra have been developed. These techniques and their implementations are reviewed in detail by Alfonso-Garcia *et al.*^[76] Briefly, such strategies include:

- 1) Wavelength sweeping either the Stokes or pump wavelength with narrow band excitation wavelengths.^[92]

- 2) Multiplex or broadband CR where one narrowband pulse and one broadband pulse are combined to create a simultaneous spectrum.
- 3) Spectral focusing wherein spectrally chirped pump and Stokes pulses are temporally swept creating complete Raman shift spectra.
- 4) Temporal sweeping of two replicas of a broadband pulse, resulting in temporal interferences.

The technique deployed for hyperspectral CR imaging is outside the scope of this review; however, in designing a clinical system the economic impact and technical difficulty of its implementation should be considered.

3.1.1.5 Spectroscopy systems for tissue characterization in neurosurgery

In recent years both SR and CR have shown potential for improving safety, accuracy and extent of resection for neurosurgical procedures. In SR, more readily available micro-optical components and more sensitive detectors have greatly increased probe efficiencies. Now, even small form-factor probes are able to detect enough SR signal for high-accuracy tissue classification at clinically relevant acquisition speeds, leading to a surge in clinical translation. For CR, the clinical adoption can be primarily traced back to breakthroughs in compact pulsed fiber-laser sources. The drastic decrease in laser size and increase in robustness has allowed the development of portable coherent Raman microscopy systems capable of being transported into the operating room on a single cart.[101]

The exploitation of the Raman effect in neurosurgery can be divided into three main system types: 1) single-point RS probes for intact tissue assessment (mainly SR systems); 2) portable Raman microscopes for rapid histopathological evaluation after tissue resection (mainly CR systems); and 3) endoscopic imagers for intact tissue histopathology and surgical guidance (SR and CR prototypes). This section provides an overview of the hardware and technical considerations required for clinical implementation of these systems.

3.1.1.5.1 Intact brain tissue interrogation using point probes

RS does not require sample preparation, and is thus able to interrogate intact and unlabeled tissue. Whilst there have been many pioneering studies in the field of neuroscience using RS, few have

bridged the gap from fundamental research to clinical utility. The first *in vivo* RS use in brain cancer patients was reported in 2015. *In vivo* implementation of CR imaging in the human brain has not been yet reported. **Table 3.1** and **Table 3.2** summarize the relevant clinical work using point probe systems; and they will also be briefly described here.

In 2007, Santos *et al.* from the Puppels group first implemented high wavenumber (HWN) RS for fiber optic brain tissue sensing.[102] By measuring only the Raman shift distant from the excitation wavelength, they showed the probe needed no distal optics to remove contaminating Raman or fluorescence background from the silica fibers. They followed this with two more studies with and without the fiber optic probe in *ex vivo* porcine brain tissue, demonstrating the ability of HWN RS to classify brain regions.[103], [104]

Another system for brain tissue sensing with an SR probe was reported by Beljebbar *et. al* in 2010, which included an analysis of Raman spectra taken from an *in vivo* mouse model of glioblastoma.[46] The probe was compact, designed professionally (SEDI, France) and acquired data in the fingerprint region.

An important advancement of Raman systems in neurosurgery was presented by Jermyn *et al.* in 2015 (Leblond's group), marking the first use of RS in living human brain tissue.[43] They succeeded in acquiring SR spectra from glioma patients in the operating room, and the system successfully discriminated normal brain tissue from cancer with 90% accuracy. Much of the system's success was enabled by a professionally designed optical probe (EmVision LLC) to maximize photon collection, and high-level data analysis procedures. Moreover, the design of the probe facilitated clinical use: it was hand-held, ergonomic, durable and had a flexible fiber. Furthermore, the probe used Neuronavigation markers on the back-end to register locations of measurements in the surgical planning suite (**Figure 3.1B**).

Since then, the same group has pushed for clinical translation of this system. They have better characterized the systems operating conditions in the operating room,[105] improved data analysis to minimize ambient light contributions[106] and compared tumor margin localization between MRI and SR.[107] The system has been commercialized (ODS Medical, **Figure 3.1B**) and is currently in clinical trials to quantify the clinical improvement of its use.

In 2017, a similar probe was presented with multiple modalities allowing it to perform SR, AF and DRS detection for increased classification accuracy.[44] This system used a different data analysis

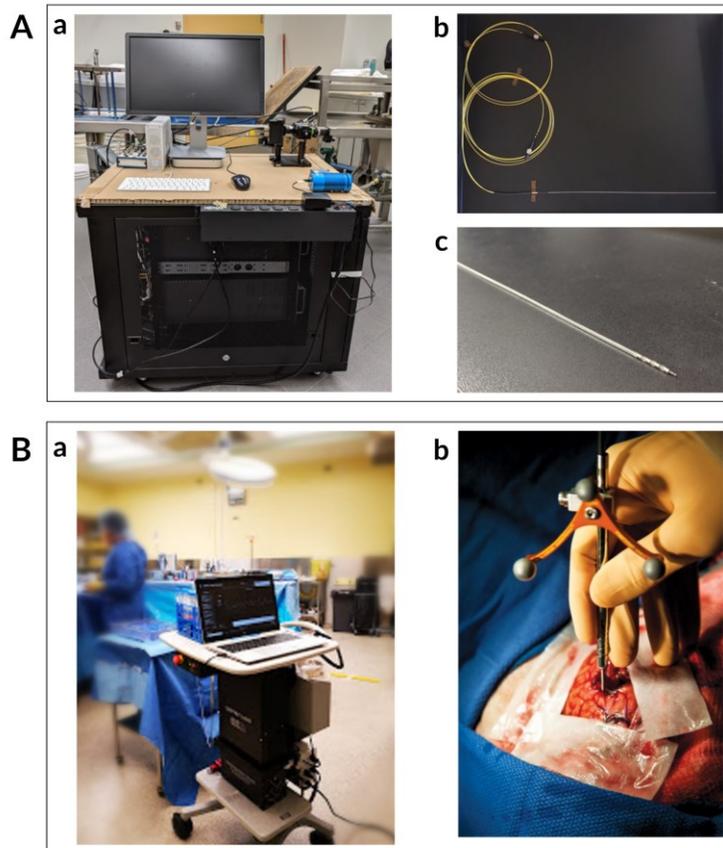


Figure 3.1: Stages of translational systems for neurosurgery. **A:** Pre-clinical, and research-oriented, system for ex vivo human surgery. Presented here is a portable CARS spectroscopy system for performing optical measurements during a DBS electrode implantation in human cadavers (Côté lab). The system consists of (a) an encased fibered laser source, an external photon counting detector and a computer for processing. The optical probe is inserted within the DBS electrode as shown in (b); however, the tip of the electrode has been cut-off to allow the probe to be in contact with the tissue (c). **B:** Clinical, and commercial, system for in-vivo human surgery. Presented here is a handheld contact fiber optic probe for SR spectroscopy, commercialized by the company ODS Medical. The system consists of (a) a 785 nm laser, a high-resolution CCD spectroscopic detector, and a computer for processing, as well as (b) a professionally designed and sterilizable contact probe for clinical use. The probe is used to interrogate live brain tissue during neurosurgery as shown in (b).

procedure and was translatable to other cancers (colon, melanoma and lung). In 2018, Desroches *et al.* followed up on the work by the Puppels group by using only the HWN region to interrogate human brain tissue intraoperatively.[63], [102] Probes were used in both *ex vivo* porcine tissue and *in vivo* human tissue, with the goal of being a proof-of-principle device for HWN tumor classification. In the HWN setup, the Raman background from silica is minimal, negating the requirement for filters at the probe tip and facilitating probe miniaturization.

The next intraoperative SR milestone was spectral acquisition from deep within the brain by Desroches *et al.* (Leblond's group) in 2019 (**Figure 3.1B**).[45] Here, the mandrin of a biopsy needle was replaced by a probe (EmVision LLC) that was optically similar to that used by Jermyn *et al.* in 2015, with the significant differences being its smaller size and its angle-facing detection.[43] The probe had some spectral discrepancies from the original, probably due to the magnesium fluoride prism used for side-reflection. Whilst this work was preliminary, it did present the first intraoperative deep brain SR measurements (both fingerprint and HWN) and opened the door for Raman-guided biopsy sampling. A variety of other probes have been implemented on *ex vivo* tissue. In 2016, Stevens *et al.* presented a probe design which used a collimated beam through an empty biopsy needle to measure signal from *ex vivo* porcine tissue.[108] This probe is unlikely to be brought into the clinic in its current form due to the free-space optics involved and the long integration times required. However, it succeeded in acquiring the signal from the low wavenumbers below 700 cm^{-1} – routinely ignored due to the background signal from silica.

In 2018, DePaoli *et al.* (Côté's group) presented a CR probe to investigate *ex vivo* primate brain tissue.[66] The system was composed of a compact fiber-based pulsed laser source (Halifax Biomedical, **Figure 3.1A**), a sensitive photon counting detector and traditional silica fiber-optic probes. Rather than being measured using a spectrometer, the spectra were encoded in time using the fast wavelength-tuning lasers. The major implication of this system was the short integration time required (10 ms for low resolution HWN spectra) given the small size of the probe. Importantly, the probe's form factor allows it to be placed within a DBS electrode hollow core for functional neurosurgery guidance (**Figure 3.1A**). However, more work is required to minimize the optical irradiance in brain tissue before iterations of this system can be used in humans. Technical information about the systems mentioned in this section can be found in **Table 3.1** & **Table 3.2**.

3.1.1.5.2 *Strengths and limitations of in situ Raman spectroscopy in neurosurgery: spontaneous vs. coherent*

There are currently a number of advantages of SR systems over CR systems for clinical spectroscopy use. Specifically, since SR is a linear process it allows several leniencies in the system design, such as:

- 1) The use of a continuous laser source rather than a pulsed laser, allowing for a smaller, less expensive system.
- 2) The use of standard silica optical fibers for transporting optical energy from the laser output to the patient. Pulsed laser systems (required for CR) are plagued by pulse-deteriorating non-linear effects occurring during their transport through a dispersive media (such as silica optical fiber).[109] Therefore, pulsed laser systems are often equipped with expensive, specialty optical fibers designed to decrease non-linear effects. This greatly increases the cost and fragility of a system.
- 3) The linear nature of SR signal generation means that the light does not need to be focused to produce the SR spectra. Resolution aside, this is an advantage as a larger excitation spot-size decreases the overall irradiance on the tissue.
- 4) SR is linearly proportional to the concentration of molecules, allowing for direct molecular quantification SRS shares this advantage. However, there are techniques to achieve linear proportionality with CARS.[95], [110]

CR is best exploited in imaging systems due to its intrinsic optical sectioning and rapid contrast at a single molecular vibrational mode. However, there may be niche uses for CR spectroscopy independent of the imaging capability. Due to the small excitation volume ($<10 \mu\text{m}$ diameter) CR spectroscopy can provide high resolution sensing, allowing the delineation of small tissue structures, such as deep brain nuclei. Furthermore, due to the optically-sectioned signal, probes can be designed so that excitation occurs far from the fiber tip, even on the other side of protective materials, without sacrificing collection efficiency (as would be the case for confocal SR). This is especially useful if the probe must be placed within a biocompatible sleeve having its own Raman signal at the interrogation wavelength.[111], [112] While the small excitation volume also means that the optical energy must be focused and confined to a small volume (and therefore limit the translational value), improved fiber lasers may decrease the required irradiance dramatically.

3.1.1.5.3 Rapid Spectroscopic blood vessel detection: an unmet clinical need

There is a clinical risk of hemorrhage when performing closed neurosurgical procedures (i.e. DBS, biopsy) since the surgeon cannot see oncoming blood vessels. Although other optical systems have been presented to fill the clinical need, Raman technologies have not yet been fully exploited for the task. Recently, a translational success was presented using intraoperative OCT for accurate blood vessel detection and size estimation from within a standard biopsy needle during neurosurgery.[68] Other optical technologies that have investigated blood vessel detection in neurosurgery are LDF and DRS; however, ability to measure blood vessel size using OCT is a significant advantage for risk assessment.[62], [113], [114]

To minimize the number of optical probes used during a single procedure, it would be ideal to have a probe capable of both blood vessel detection during needle descent and tumor margin detection. Such a probe could be either Raman or OCT, or size permitting, a multimodal combination.

3.1.1.5.4 Rapid and portable Raman microscopes for operating room histopathology

Histology is time-consuming, requiring fixation, sectioning, and staining of freshly excised tissue. Furthermore, stains require interpretation which can prove challenging or ambiguous even to trained pathologists. Raman technologies, however, provide molecular information with minimal tissue preparation. This would be particularly beneficial in cases where delayed diagnosis could lead to possible repeat surgery because of residual cancer tissue. In this section we will overview the work that has been done towards vibrational imaging systems for intraoperative *ex vivo* neuropathology, typically with the objective of providing rapid point-of-care pathology information during surgical interventions.

SR microscopes have proven fundamental to dissecting Raman differences in neuropathology;[115]–[118] however, they are traditionally too slow to be used intraoperatively, requiring hours to provide images of tissue slides at microscopic resolution. On the other hand, CR systems built upon the molecular knowledge acquired using SR have shown potential for clinical translation.

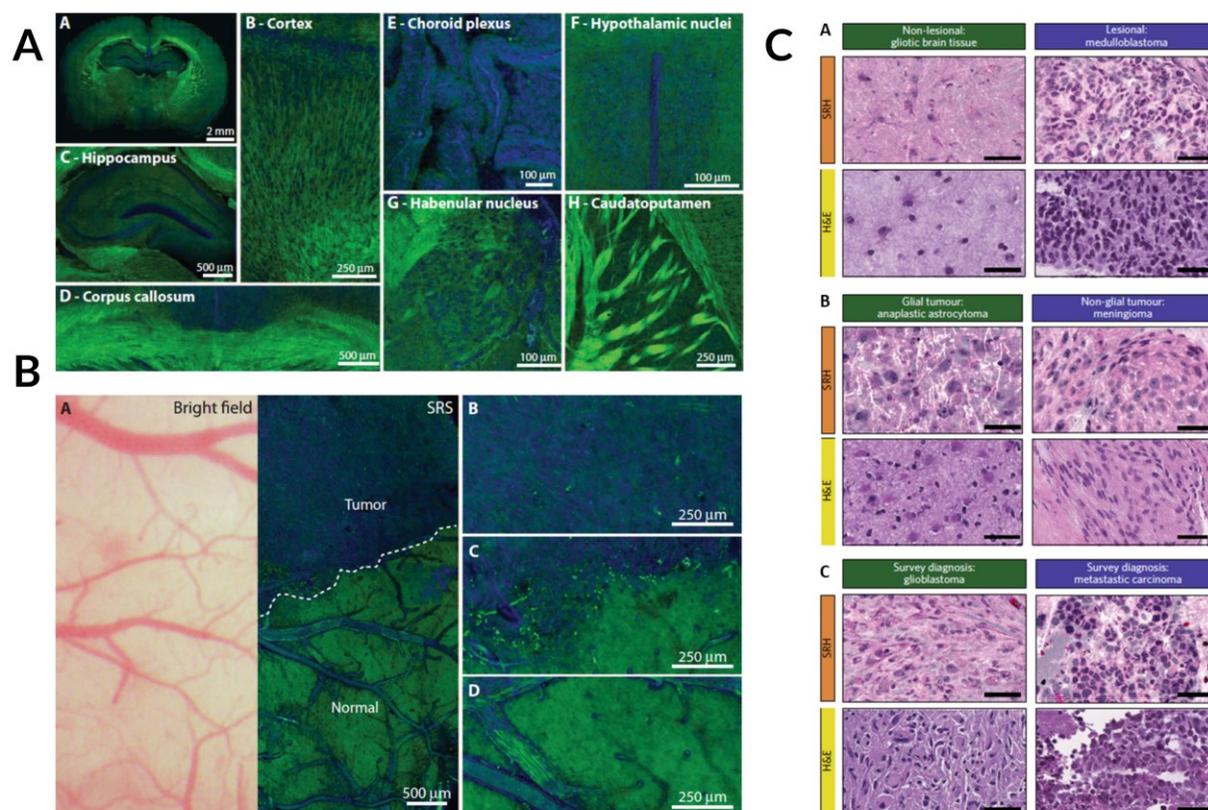


Figure 3.2: SRS contrast in brain tissue. **A:** Epi-SRS images of fresh brain slices from normal mice brain in various brain regions. Lipids are shown in green and proteins in blue. Referring to the inset sub-figure labelling: 2-mm coronal slice (A), cortex (B), hippocampus (C), corpus callosum (D), choroid plexus (E), hypothalamic nuclei (F), habenular nucleus (G), and caudato-putamen (H). From Ji et al., 2013. **B:** Brightfield and SRS imaging through cranial window, 24 days after implantation of human GBM xenografts, for comparison of information. Referring to the inset sub-figure labelling: (A) Same FOV bright-field and SRS image of xenograft boundary.

Brightfield appears normal, whereas SRS microscopy within the same FOV demonstrates distinctions between normal and tumor-infiltrated areas. (B to D) Higher-magnification views of tumor (B), at the tumor-brain interface (C), and within normal brain (D). From Ji et al., 2013.

Reprinted with permission from AAAS. **C:** Comparative examples of processed SRH and H&E images of gliotic brain tissue, medulloblastoma, anaplastic astrocytoma, meningioma, glioblastoma and metastatic carcinoma. These images were used in a web-based survey to compare diagnostic outcomes using the two histological methods. From Orringer et al., 2017.

Reprinted by permission from Springer: Nature.

The first use of rapid CARS imaging for healthy and cancerous mouse brain tissue delineation was

by Evans *et al.* in 2007.[119] This was implemented *ex vivo* on an orthotopic human astrocytoma mouse model and the tumor boundaries were defined by the reduced CH₂ signal in the tumor regions. Since then, there has been a considerable push for CARS-based histology. In 2014 Uckermann *et al.* demonstrated a reduction in lipid signal in infiltrative tumor regions in an orthotopic glioblastoma and brain metastasis mouse model using CARS.[120] By combining CARS with modalities such as two-photon excited fluorescence (TPEF) and second harmonic generation (SHG), detailed images of tissue with structures such as extracellular matrix, blood vessels, and cell bodies could be created. Other groups have further demonstrated the ability of intrinsic TPEF and SHG to aid CARS in brain cancer histology.[121] Galli *et al.* also performed multimodal imaging on excised human tissue samples after 5-aminolaevulinic acid (5-ALA) was preoperatively administered and showed that it did not interfere with the CARS signal.[122] In 2019, the group used the same multimodal approach in an endoscopic setup and demonstrated that the findings were comparable to those *in situ*.[123]

The main issue with using CARS for histology applications is the low vibrational contrast from proteins, usually represented by CH₃ contrast. This somewhat limits the ability of CARS images to be directly compared to the gold standard of hematoxylin and eosin (H&E) staining for pathology.[124] Stimulated Raman scattering, on the other hand, does not suffer from this shortcoming. In 2012, Freudiger *et al.* first used SRS to create images with H&E type information using the vibrational contrast from only CH₂ and CH₃ bonds.[124], [125] A year later, the group showed a high correlation between SRS histology and H&E staining ($\kappa = 0.98$) for glioma detection in mouse brain tissue.[99] Using an backward-illumination and detection SRS microscope, the group guided the resection of a mouse brain tumor *in vivo* (**Figure 3.2**).[99], [126]. In 2015, the group reported continued progress in using SRS histology to accurately detect and automatically classify tumor infiltrated tissue sections with high accuracy in human brain.[100] This suggested the feasibility of pathologist-free interpretation of tumor margins for rapid-feedback in the operating room.

In 2016, Lu *et al.* took the translatability a step further by analyzing fresh human samples, showing that some additional discerning features seen on SRS images were lost in the tissue preparation phase of H&E staining.[127] A particularly innovative aspect of this work is that the images are freely available to help improve diagnostic training in the future.[128] Made possible by advances in portable and robust fiber laser systems;[129] the first true fruition of CR *in situ* potential was

presented by Orringer *et al.* in 2017, wherein the group reported a portable clinical SRS system for intraoperative *ex vivo* neuropathology (**Figure 3.3A**).[101] Using this system, freshly resected tissue sections were compressed and imaged within the operating room. By taking several small field of view (FOV) images, they created interpretable SRS histology mosaics on the order of 2.5 minutes. In terms of output, the Raman information was used to digitally recreate H&E type staining (**Figure 3.2C**), and to perform automatic tissue classification using these images. The optical characteristics of the system are presented in **Table 3.3 & Table 3.4**. Furthermore, the system has remained in use for over a year within the operating room without problems or realignment, which speaks to its robustness. This is a point not often mentioned in optical reports but imperative in clinical designs. In 2018, the group reported the use of this system for *in situ* pediatric brain tumor classification with 100% accuracy.[130]

Finally, in 2018, Bae *et al.* presented an epi-illumination and detection hyperspectral SRS system for the subtyping of glioblastomas using HWN spectra.[131] Technical information about the systems mentioned in this section can be found in **Table 3.3 & Table 3.4**.

3.1.1.5.5 Towards Raman endoscopes for label-free imaging of intact tissue in neurosurgery

SR systems have dominated clinical implementation of fiber-delivered spectroscopy, while CR systems are the more popular option for pre-clinical biological imaging applications. This creates a crossroads at imaging endoscopy for the two types of modalities. On the one hand, SR systems used for imaging (i.e. moving beyond single-point) are currently too slow for *in vivo* imaging as the signal is intrinsically weak and dispersed spectrally. On the other, CR systems traditionally require bulky lasers and complicated optical transport methods for the high-peak power pulses and provide restricted spectral information. However, there have been considerable technological advances in the past decade for both system types showing that vibrational endoscopy in neurosurgery is close to a reality.

To date, there are few SR imaging endoscopes reported to date and none have been deployed for neurosurgical improvement. There have, however, been attempts at handheld systems for mesoscopic Raman imaging. St-Arnaud *et al.* (Leblond's group) presented two iterations of a macroscopic wide-field Raman imaging system with ~1 cm field of view and <400 microns resolution using a multicore imaging fiber for image transport and a tunable filter in the detection path for temporally-encoded Raman spectra (**Figure 3.3C**).[132], [133] While the system required

~1 minute integration times and high optical power, due to the widefield illumination irradiance levels were actually quite low. In the future, specific Raman bands could be selected for imaging to decrease imaging time.

CARS endoscopes have been under investigation since 2006, when Légaré *et al.* presented backwards imaging of polystyrene beads using a single-mode fiber for both illumination and collection.[134] Subsequently, other studies have improved our understanding of the inherent limitations of traditional silica-based CARS endoscopes.[135], [136] Due to these limitations, specialty fiber optics for pulse delivery with reduced dispersion and background Raman signal have been investigated.[137]–[139]

However, CARS endoscopes with the possibility for clinical translation have only recently been reported. While none of the systems have been used *in vivo* or in brain tissue, these promising candidates for neurosurgical use will be included here. In 2017, Lukic *et al.* presented a multicore imaging fiber system which allowed for multimodal CR imaging with no moving parts at the 8 mm outer diameter probe tip.[140] Using this system they imaged a skin tissue sample with 300 μm field of view and a 10 s acquisition time. While the probe is quite large for neurosurgery, considerable downsizing could be possible with micro-optical components. In 2018, Lombardini *et al.* reported a high performance CARS endoscope with $<1 \mu\text{m}$ resolution, capable of producing CARS images of a $310 \times 310 \mu\text{m}$ FOV in only 0.8s (**Figure 3.3D**). [141] The outer diameter of this probe is 4 mm and images were presented on fresh colon tissue. This system is also capable of variable FOVs and multimodal imaging. The high performance achieved is mainly due to the sophisticated design, using specialty optical fibers (double clad, Kagome lattice) and a precision-spliced microlens. While this may seem fragile for the clinic, it is part of the tradeoff for high performance imaging systems. Finally, also in 2018, Zirak *et al.* presented a 2.2 mm OD rigid CARS endoscope (187 mm in length) for neurosurgery applications. The endoscope was shown to be capable of high resolution, fast CARS imaging with the smallest OD to date.[142] The technology is enabled by recent advances in GRIN lens technology, whereby long versions of the image-conserving fibers (previously used for *in vivo* CARS endoscopy in mice[93]) are now capable of being manufactured. That said, the system still uses large free-space lasers and an optic table for alignment. Therefore, while it is promising, some engineering is required to make the system fibered and ready for the clinic.

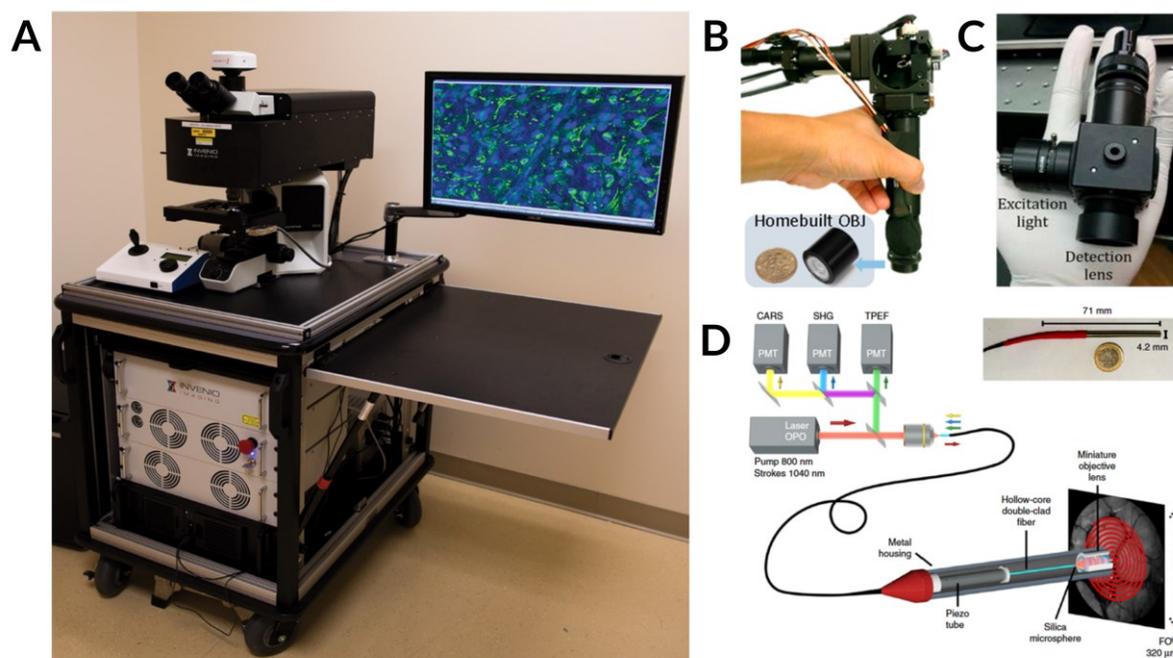


Figure 3.3: Translational Raman imaging systems. **A:** On cart SRS microscope for intraoperative imaging of freshly resected brain tissue. From Orringer et al., 2017.[1] Reprinted by permission from Springer: Nature. **B:** Fiber-delivered, Handheld SRS microscope. From Liao et al., 2018.[1] Reprinted by permission from American Chemical Society. **C:** Handheld widefield SR imager with large FOV. From St-Arnaud et al., 2018. Reprinted by permission from Wiley. **D:** Multimodal CARS, TPEF, SHG endoscope used to image human colon. From Lombardini et al., 2018. Reprinted by permission from Springer: Nature.

In 2010, Saar *et al.* presented the first report of fast epi-detected SRS with acquisition times of ~ 100 ms using 50 mW for both the pump and Stokes wavelengths.[126] However, the detection scheme in the report used a 10×10 mm photodiode with a hole drilled in the center, through which the excitation lasers were focused. A year later Saar *et al.* presented a scanning-fiber-endoscope version of the system using ~ 130 mW total power for excitation and the same detection apparatus; *in vivo* work has not been presented since using the device.[143] In 2018, building on earlier work in delay-line tuning, Liao *et al.* presented a handheld hyperspectral SRS microscope capable of HWN spectroscopic images (15cm^{-1} spectral resolution) on the order of 3s (**Figure 3.3B**).[144], [145] Using this system, the team imaged sections of healthy and cancerous canine brain tissue, but did not go into much detail on the ability to distinguish the two samples. Technical information about the systems mentioned in this section can be found in **Table 3.3** & **Table 3.4**.

3.1.1.5.6 CARS or SRS endoscopy

Stimulated Raman spectroscopy has improved nuclear contrast in comparison to CARS due to a reduced non-resonant background, allowing faster imaging speeds of molecules at low concentrations.[95], [125], [146] However, successful systems imaging in the backward (epi) direction using SRS are scarce in comparison. In microscopy, this is likely because CARS can be more easily incorporated into traditional laser scanning microscopes used for TPEF and SHG. In clinical applications, however, the main hurdle for SRS is the increased complexity of signal detection, made even more complicated with the push for portable fibered lasers. Gottschall *et al.* provide a good resource for understanding the benefits of each CR modality in their review on advances in laser concepts for multiplex CARS.[147]

Another factor which can play a considerable role in clinical applications is the operability of the system under ambient lighting. SRS has an advantage here for two reasons: 1) the modulated signal can easily be distinguished from background contributions and 2) the wavelength of detection is further into the near-infrared region. Although traditionally considered an advantage of CARS in microscopy, the blue-shifted signal generation is a burden in the clinic as the detected signal is often near the visible region where there are strong surgical lighting contributions. This could be circumvented by pushing CARS sources further into the NIR region.

3.1.1.6 Optical exposure to brain tissue

There are limits to the amount of optical irradiance which can be introduced into a biological system without causing serious cellular damage. The maximum permissible exposure (MPE) of continuous-wave (CW) optical radiation for tissues such as the retina and the skin can be calculated from international standards.[148], [149] There is also an inherent danger of laser light accidentally being shone directly into an eye at any point during a laser's operation. While this is a serious challenge in designing clinical laser instruments for ethics approval and eye-safety, here we speak more about the dangers of deliberate laser-tissue interactions, and assume that the proper safety eye-wear, or other appropriate risk mitigation strategies (e.g. laser activation only upon tissue contact with an imaging probe), are being used in the operating room.

Although the international standards provide a convenient calculation for the MPE of CW radiation, they should not necessarily be used as a guideline for Raman systems investigating brain tissue. The optical and thermal properties of skin and retina differ considerably from brain tissue

and the higher water content of the brain results in a lower conversion of photon energy to thermal energy.[150]–[152] Moreover, the standards themselves are not designed for deliberate laser exposure during medical procedures.[148], [149] To make matters more complicated, brain sensitivity to thermal damage is somewhat unclear. Reports of minor local temperature changes of only 2°C have been shown to cause thermal damage to metabolically active brain cells.[152] However, it has also been shown that the awake animal brain naturally fluctuates in temperature within this range of 2°C.[153] Furthermore, in terms of photothermal damage at the cellular level it has been shown that injuries are reversible for temperatures that have increases of 6°C.[154]

The photo-induced effects of coherent Raman systems are particularly complicated due to: 1) high power density, 2) focal point scanning, and 3) non-linear damage such as photochemical ablation and optical breakdown.[155]–[157] Due to short pixel dwell times in rapid focal point scanning, instantaneous heating through linear absorption is often considered negligible in comparison to non-linear damage for multi-photon systems.[155], [156] However, continuous scanning of the same FOV can produce a volumetric heating effect that must be accounted for.[114], [158] In 2001, Hopt *et al.* described a general formula for the tissue damage rate, D , which was proportional to the optical intensity raised to the n th power (P^n), the repetition rate (f_{rep}), and the pulse width (τ_{pulse}) of the laser. Using this information, Gotschall *et al.* explain an optimal CR system which can maximize signal while minimizing damage.[155] Such a system for clinical CR would operate in the NIR (pump & stokes = 1000 - 1500nm), have a relatively low repetition rate near 1 MHz and use pulsewidths between 1 - 10 ps, operating at about 30 mW total average power. [155]

3.1.1.7 Data analysis for spectroscopic information

Working with Raman data presents multiple challenges: Raman signal is intrinsically weak, and complicating factors such as excess blood, surgical lighting, and device manipulation can exacerbate this. Furthermore, living tissues are complex and dynamic systems, composed of thousands of interacting molecules that are heavily influenced by external factors.[159]–[161] Competencies at the intersection of signal processing (spectral, image-based or a combination of both), data mining and machine learning (ML) are essential to design cutting-edge biomedical Raman spectroscopy systems.

3.1.1.7.1 Spectra data processing

Raw Raman acquisitions are characterized by low Raman signal, high amounts of shot noise and intense signal from background sources (e.g. AF and ambient light in SR, coherent background in CARS).[105] Signal processing aims to maximize the Raman component of the acquired signal while minimizing the contribution of these other processes. As an example, in SR a complete signal processing pipeline generally includes: truncation of the signal to the desired spectral range, correction for ambient light, cosmic ray removal, correction of the spectra for the system response with a standard measurement, background removal, smoothing of the spectra to remove high-frequency signal associated with shot noise, normalization and data quality assessment (**Figure 3.4**).[160]

Following cosmic ray removal and correction for system response, the most important remaining signal contribution in SR is AF, a spontaneous process resulting from the emission of light from endogenous tissue fluorophores including elastin, tryptophan and nicotinamide adenine dinucleotide (NAD).[162] Even with proper hardware, AF intensity can be orders of magnitude higher than that of Raman scattering.[163] AF results in a broad spectrum that can underlie the narrow Raman peaks, making baseline estimation a critical step in the signal processing routine. Many techniques exist to mathematically estimate and remove the background in SR, leveraging the smooth and predictable decay of its contribution throughout the spectral range. Polynomial fitting of the spectra is the most widely used technique, but heavily relies on expert knowledge to select the proper parameters and avoid over- or under-fitting the signal.[160], [164] In order to create completely automated routines, more recent algorithms have surfaced that rely on iterative fitting of the signal with sophisticated cost functions ensuring an improved fit while minimizing expert intervention.[165]–[168] Nevertheless, in biological experiments the background-generating processes are not always clearly identified, resulting in a correction that is more based on spectral morphology than exact comprehension of the underlying phenomenon. Some authors even argue against this step to avoid altering the spectral shape in unpredictable ways.[169]

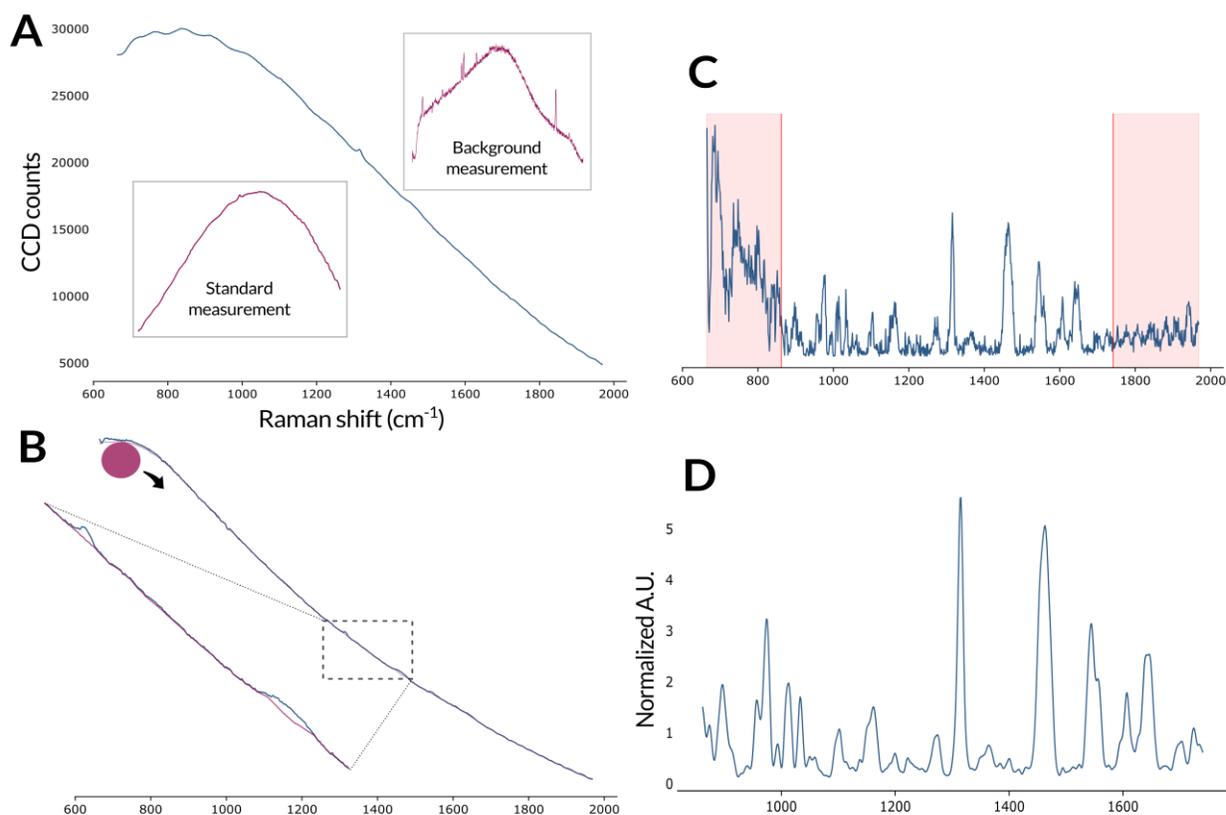


Figure 3.4: Spectral processing of raw Raman signal in biological tissue. **A:** Raw Raman measurement (blue). During the acquisition, a background measurement is also recorded with the laser turned off to correct for ambient light. After each experiment, an acquisition on a Raman Standard with known Raman response is used to correct for artefacts from the acquisition system. **B:** Baseline correction. A curve fitting algorithm (pictured here: Rolling Ball algorithm) is used to estimate the shape of the baseline signal, which mainly consists in tissue auto-fluorescence. This baseline curve is then subtracted from the acquired signal. **C:** Truncation to desired spectral range. Spectral regions with poor Raman information or exhibiting artefacts from the experimental design (e.g. silicate substrate) or correction algorithm are removed. **D:** Smoothing and Normalization. High-frequency noise is removed from the signal and the spectra are expressed in normalized units so that they can be compared across samples and experiments.

A.U.: Arbitrary units.

3.1.1.7.2 Data analysis methods

Whilst Raman spectroscopy has historically relied on the visual assessment of the spectra, the complexity of the signal acquired in living tissue has necessitated the reliance on automated algorithms.[170] Most clinical applications of Raman spectroscopy for neurosurgery, rely on three types of analytical tasks: supervised ML (classification), spectral imaging, and biomolecular interpretation of spectral features.

3.1.1.7.2.1 Supervised machine learning

Supervised learning consists of training an algorithm; for example, to recognize tissue phenotype from its spectral signature. A data matrix $\mathbf{X}_{n \times p}$ containing n spectra represented each by p variables or features (e.g. the signal intensity value for each wavenumber) is associated with a vector of labels y (e.g. the tissue diagnosis from the neuropathologist) (**Figure 3.5A**). The algorithm or function f for which $f(\mathbf{X}) = \hat{y}$ is optimized to minimize the loss function $L(y, \hat{y})$, where \hat{y} is the model-predicted tissue label. The result of the loss function is called the training error. The testing error is calculated from predictions on new data not used for training. Common supervised learning algorithms are described in **Table 3.5**.

Feature engineering is the transformation of the processed signal into a set of variables that will be used as representation for the learning task. It ranges from selecting a subset of all available intensities associated with different wavenumbers, to generating new variables and performing complex mathematical transformations of the data to unveil properties not necessarily conveyed by the original spectrum. In Raman spectroscopy, this transformation is critical because of two properties of the data: high-dimensionality and sparsity of the feature space.[171] High dimensionality refers to the large number of available features (e.g. between 500 and 1,000 spectral bands in SR) needed to describe a single observation. Feature sparsity means that most of this information will be unhelpful in discriminating different tissue phenotypes: all tissues are composed in majority of the same organic molecular compounds and many of the molecular markers are redundant across the spectral range.

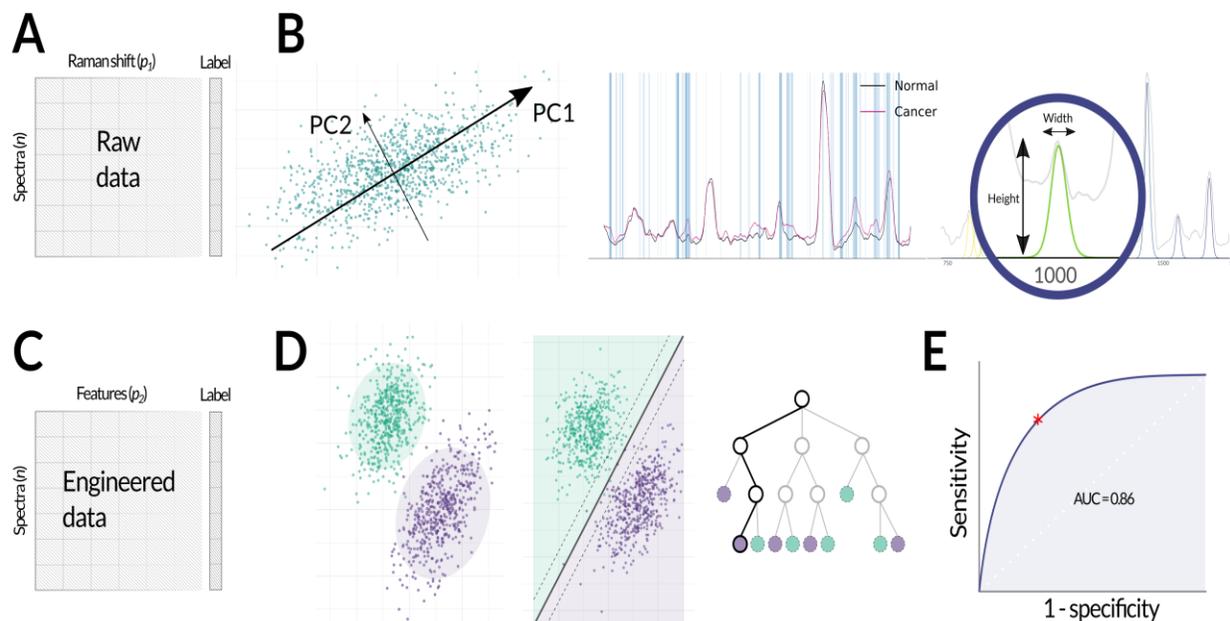


Figure 3.5: Machine learning for Raman spectroscopy. **A:** The initial data matrix is composed of n spectra and p_1 variables, where each variable represents an intensity at a particular wavelength.

A standard, processed Raman spectrum can contain between 500 and 1000 variables. Each spectrum is associated with a label (e.g. *High-grade glioma* vs. *normal tissue*). **B:** Feature engineering algorithms. These algorithms are designed to change the representation of the initial spectra matrix into one that will enhance performances of the ML predictions. From left to right: PCA, feature filtering and feature extraction (peak fitting). Each method is described in the main text. **C:** The engineered data matrix. Each spectrum now contains p_2 variables, which are the results of the previously applied feature engineering methods. These variables could be PC scores, intensity at specific Raman shifts, peak-fitted peak height, etc. **D:** Supervised learning algorithms. Different mathematical function can be trained to separate the spectra with distinct labels based on the values of their variables. From left to right: LDA, SVM and Decision tree. **E:** Testing of the ML model. The different ML pipelines (including feature engineering) are trained on a subset of the data and tested on an independent subset. The predictions of the models are compared to the true labels of the testing set, and prediction performance is evaluated. A ROC curve can be drawn to estimate the performance of the prediction at varying levels of sensitivity and specificity. The best performing threshold is the point on the curve closest to the left upper corner (red star). PC: Principal Component, PCA: Principal Component Analysis, AUC: Area-under-the-curve.

In the neurosurgical literature, the most common feature engineering algorithm is principal component analysis (PCA).[46], [103], [116], [118], [169], [172]–[180] PCA iteratively finds the orthogonal vectors (or principal components, PC) that maximize the variance in the dataset, and then stores the projection of the data point upon each PC (**Figure 3.5B**). The PCs are ranked by their eigenvalues, or amount of variance explained; therefore, most of the variance in the dataset can be expressed in the first PCs. Subsequent PCs can be discarded, resulting in a compressed dataset. The number of retained PCs ranges between 2–40 depending on either pre-defined criteria such as amount of explained variance, [46], [104], [116], [175]–[177], [179], [180] or post-hoc criteria e.g. selecting PCs that could better differentiate the different tissue types.[118], [172]–[174] The popularity of PCA in Raman spectroscopy can be attributed to its unsupervised nature: as it is agnostic to labels, it is considered unbiased.[48] Moreover, most authors report that over 99% of the variance of their dataset is expressed in the first 2–40 PCs[46], [104], [115], [175]–[177], [179], [180] and the orthogonality of the extracted features can improve the efficiency of classical multivariate linear models.[181]

Another feature engineering technique involves selecting a small subset of the best features and discarding all others. These features can be identified by: 1) simultaneously assessing their individual correlation with the outcomes (called filter methods),[182]–[185] 2) iteratively evaluating changes in ML performances when excluding/including each feature into the model (wrapper methods),[186] or 3) adding penalization terms to the optimization of the training algorithm so that some of the features contribution are reduced to zero (embedded methods).[187] The features are selected either directly from the processed spectra, or from previously extracted features, as part of a two-step feature engineering pipeline.[172]–[174], [177], [188] Most studies involving brain tissue rely on filter methods to identify important features,[189], [190] and in some cases, features are selected manually.[188], [191], [192] In proteomics and genomics studies, performance-based and optimization-embedded have proven superior to filter methods in recovering truly important features in sparse, high-dimensional datasets.[193], [194] Embedded methods have recently started to emerge in the biomedical Raman literature to identify crucial features, but not as part of a supervised learning task.[195]

The last feature engineering strategy is the generation of new features from the existing data. For example, the height ratio of two peaks can carry important information such as summarizing the lipid-to-protein content of a sample. Peak ratios have proven useful in discriminating white and

grey matter in the brain, and in differentiating between normal, necrotic and malignant tissue. The shape of the Raman spectrum can also be used as marker of malignancy.[188], [195] In their study, Stables *et al.* selected eight target bands, for each of which they calculated the centroid (weighted mean of the signal in the defined region), skew (asymmetry of the intensity values) and kurtosis (prominence of certain intensity values from the rest of the bands). This way, a sequence of 10-30 contiguous variables within a Raman tissue band can be represented by two or three features or parameters, resulting in enhanced ML performances (**Figure 3.5B**). While this approach to feature extraction is extensively used in the Raman literature,[196] it is rarely used in neurosurgery-related research.[48], [197] Opponents of band fitting for biological Raman signals argue that the selection of target bands cannot be reliable in the case of low SNR signal exhibiting a high number of potential peaks, and that our knowledge of Raman generating processes in tissue is not strong enough to limit our analysis to a few critical bands.[48], [197] Over time, more authors may begin to incorporate band-fitting routines to the analytical pipeline, as it is an effective way to summarize the vibrational profile of a tissue in a biochemically meaningful manner, while allowing for easy statistical manipulation and even further feature engineering.[195]

3.1.1.7.2.2 Classification

The classification tasks comprise of the selection, training and evaluation of a ML model that will map the engineered feature matrix to a vector of observed tissue classes (**Figure 3.5D-E**). Linear discriminant analysis (LDA) is the most tried and tested Raman classification model in RS. LDA assumes that all observations with the same label originate from multivariate normal distributions with equal covariance and assigns new observations to the label with highest likelihood. Because it is vulnerable to highly correlated features, LDA is often preceded by PCA. Despite its widespread adoption, LDA relies on a set of assumptions that are not respected by Raman data: multivariate normality of the distributions, independence of the predictors, homoscedasticity (homogeneity of covariance across labels) and few outliers.[198] Furthermore, because of the often inconsistent SNR characteristics of biological acquisitions, researchers need algorithms that are robust, i.e. that are not overly sensitive to noise in the data. The complexity of biological Raman data and easy access to ML libraries have motivated the reliance on more flexible models for this problem.

The support vector machine (SVM) learning model finds the hyperplane that optimizes the distance between itself and the closest point of each class it tries to separate.[199] By using kernels, SVM

can adapt to a non-linear feature space with minimal computational cost. Authors who have used SVM were able to separate up to seven tissue classes with acceptable accuracy using SR signal from *ex vivo* brain tissue.[175], [179], [188], [200] However, extensive feature engineering was necessary in all cases. In the anticipation of the translation to clinical practice, other types of models are being explored to further increase predictive performances of Raman-based systems. Decision trees are non-parametric models where each internal node represents a decision function based on an input variable. Ensemble methods such as bagging and boosting, which work by aggregating multiple decision trees trained on different subsets of the data, can significantly improve the performances of decision trees[201]. Boosted decision trees have shown success in classifying between normal tissue and glioma without using any prior feature engineering.[43], [44] Artificial neural networks (ANN) have received the most attention in ML research over the last few years due to their potential deep structure and their ability to adapt to virtually any possible function.[202] In neurosurgical applications, Jermyn *et al.* (Leblond's group) showed how a simple ANN could reduce interference with surgical ambient light, while performance of other models was strongly affected in such conditions.[106] Deep neural network architectures such as convolutional neural networks (CNN) have shown great promises in biospectroscopy, with the additional benefit of being less dependent on spectral pre-processing.[203]–[205], [206, p.], [207]–[209] Data requirements to train and optimize such models are high because of the millions of parameters they contain;[210]–[212] nevertheless, open access to large Raman datasets and strategies such as transfer learning and novel data augmentation methods (such as the simulation of Raman spectra for DRS analysis[213]) will make their adoption possible for biomedical applications in a close future.[214]

3.1.1.7.2.3 Single-band to hyperspectral imaging

Raman measurements acquired at regular intervals over a sample can be assembled into a Raman image, where each pixel contains a Raman spectrum. Raman imaging can solve several limitations of point probe systems as they can better resolve the heterogeneity of a sample and can work across scales, from microscopic imaging to wide-field, macroscopic imaging. Furthermore, the contrast in these hyperspectral images can be tuned to contain varying levels of molecular significance, depending on the way the spectra are processed and analyzed.

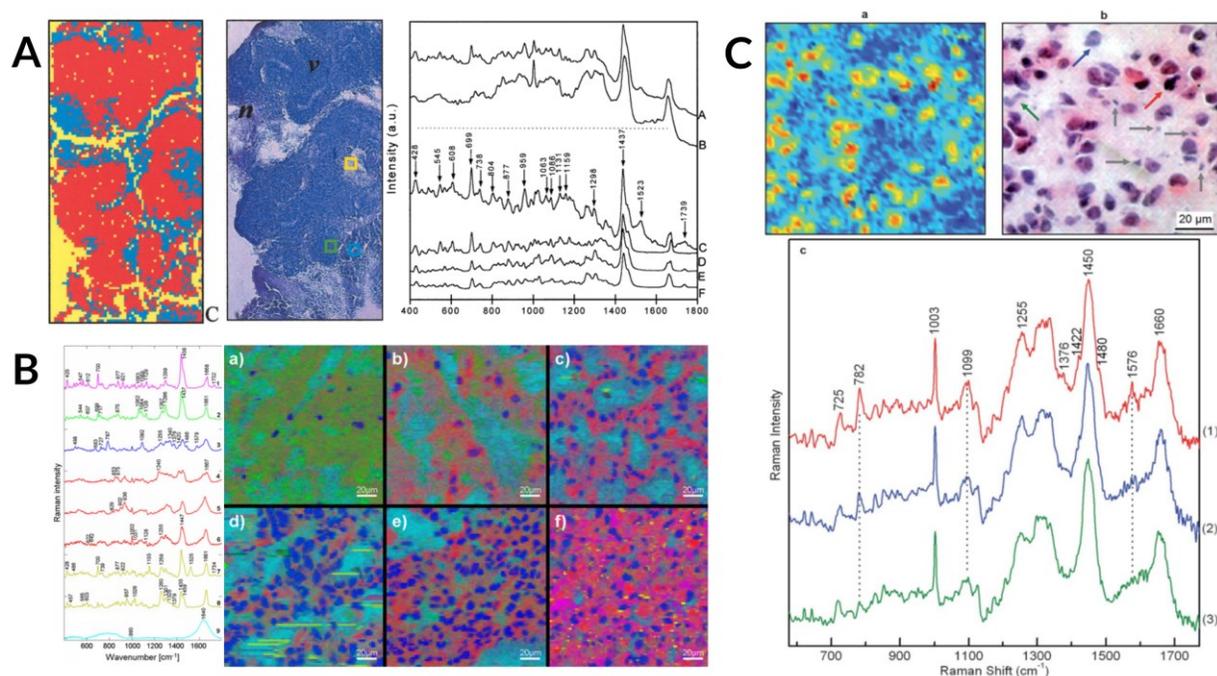


Figure 3.6: Spectral unmixing and clustering-based approaches to Raman imaging, where every pixel consists in an entire processed Raman spectrum. **A:** ML -based approach. A supervised ML model was used to assign each spectrum acquired in glioblastoma tissue to one of two classes: vital tissue and necrotic. Left: the Raman image, color-coded with the predicted class for each pixel (red: vital, blue: necrotic, yellow: background). Center: H&E image of the sample (*v*: vital, *n*: necrotic). Right: The averaged spectra of the *necrotic* (A) and *vital* (B) samples, with the difference spectrum (C) compared to cholesterol (D), cholesterol oleate (E) and cholesterol linoleate (F). From Koljenovic et al., 2002. Reproduced by permission from Springer: Nature. **B:** Endmember-based approach. The N-FINDR unmixing algorithm was used to identify a pre-specified number of endmember spectra from a dataset acquired in glioma tissue. Each spectrum in the dataset is then expressed as a linear combination of these endmembers. The endmembers were assigned to cholesterol ester (1, magenta), phospholipids (2, green), DNA (3, blue), proteins (4--6, red), beta-carotene (7, yellow), unsaturated fatty acids (8, yellow) and phosphate buffer solution (9, cyan). Each pixel in the Raman image is then colored based on the relative abundance of each endmember in its spectrum. From Bergner et al., 2013. Reproduced by permission from Springer: Nature. **C:** Endmember-based approach. VCA was used to find three endmember spectra in a dataset acquired in glioblastoma samples. Each endmember was assigned to a RGB channel (spectra 1, 2, and 3). The relative abundance of each endmember

in every pixel is used to color-code the Raman image (a). The H&E image (b) is provided as a comparison. From Krafft et al., 2012. Reproduced by permission from the Royal Society of Chemistry. VCA: Vertex Component Analysis.

Unsupervised learning is a family of ML algorithms that does not rely on the pre-labeling of each observation; instead, they serve to unveil hidden patterns in the data. An example of unsupervised learning is clustering: the dataset is divided in several groups or clusters where the similarities between spectra in a cluster is maximized compared to the dissimilarities with spectra from other clusters. This method has the advantage of accounting for the entire spectral information available and not just one or two defined bands. Early adopters of this technique were Koljenovic *et al.*, who used PCA and clustering to assign each pixel to one of 70–72 distinct clusters.[177] A supervised model was then trained to map a cluster to either necrotic or vital tissue, on which the final contrast of the images was based. Other authors have used clustering, with or without prior engineering, to assign each pixel to a specific cluster and generate the contrast[176], [215], [216] (**Figure 3.6A**).

When more than two clusters (or colors) are used, clustering-based images can become challenging to interpret. Recent approaches attempt to overcome this limitation and create images more amenable to visual interpretation. As SR signal is a linear combination of the individual Raman signals of every molecular compound in a tissue, recovering the coefficients of this linear combination can uncover the relative quantity of important molecules from the tissue Raman spectra. This approach has been referred to as hyperspectral unmixing. The different methods for linear unmixing are based on three elements: the selection of the basic, reference substances or endmembers, their spectral signature and their relative abundance. In brain tissue, vertex component analysis (VCA) and N-FINDR are two such algorithms. Both are unsupervised, and mostly differ in their definition of endmember spectra. In N-FINDR, reference spectra from pure molecular substances are set by the user, and endmembers are set as the spectra with highest correlation to those reference substances. In VCA, the endmembers spectral signatures are mathematically identified based on the variations in the dataset. Importantly, both methods assume that signal from pure substances are present in the data, and that the molecular compounds are independent.[217], [218] Applied to brain tissue, VCA and NFINDR were used to create pseudo-color images with RGB channels, each channel representing the relative abundance of one of three endmembers[169] (**Figure 3.5C**). The number of distinct channels used to create images varied from 3 to 9 in some studies, with higher number yielding images able to discriminate between more

different molecular components[115], [116] (**Figure 3.6B**). With this approach, researchers were able to generate metrics such as nucleic acid content, lipid content and lipid-to-protein ratio which correlated with malignancy in astrocytoma samples.[116] An issue with the spectral unmixing of Raman pixels is the difficulty in assessing the reliability and variance in both the endmembers definition and the relative abundance estimation. While spectral unmixing has the advantage of harnessing information from the entirety of the Raman spectral range, more recent work settled for extraction of specific, pre-defined spectral features such as band ratios and single-band intensities to represent independent molecular information.[117], [219] Nonetheless, it demonstrated a 90% pixel-to-pixel classification accuracy between white matter, grey matter and pathological brain tissue (glioblastoma, necrosis or infiltrating cancer).[117]

Raman images can help pathologists visualize molecular information not present in H&E stained samples, but they could also be amenable to ML classification algorithms. This would considerably reduce time-to-diagnosis during neurosurgeries, where the aggressiveness of the resection can be dictated by a preliminary diagnosis provided by a neuropathologist from a frozen tissue sample. Recently, Hollon *et al.* have used a CNN to classify Raman-based images of freshly excised brain tissue into one of 13 histologic categories[220]. In a prospective, multicenter study, they demonstrated a classification accuracy of 94.6%, while automated classification based on conventional H&E staining was 93.9%.

3.1.1.7.2.4 Biomolecular identification of spectral features

Despite inferring a diagnosis for a tissue sample, Raman spectra can be interrogated to describe the biochemical content of a sample. In a static system composed of a single molecular compound acquired in perfect experimental conditions, the Raman features of the distinct vibrational modes are directly observable, and subtleties such as a slight peak shift or peak widening can be used to derive changes in the system.[221] In biological tissue, computational strategies are needed to deconstruct the more complex Raman spectra.[203], [222]

A widespread approach is the comparison of specific spectral features (e.g. peak heights, band intensities, band-to-band ratios) between the averaged spectra of each tissue phenotypes (**Figure 3.7A**). These bands are either recovered from the difference spectra between two tissue phenotypes, or by analyzing the relative importance of each Raman bands after applying PCA. In brain tissue, Raman signal originates predominantly from nucleic acids (782, 829 and 1339 cm^{-1}), lipids (1063,

1086, 1131, 1268, 1300, 1441, 1659, 1670 and 1739 cm^{-1}), Amide I (1659 cm^{-1}) and III (1268 and 1300 cm^{-1}), and amino acids such as tyrosine and proline (829, 852 and 877 cm^{-1}), tryptophan (877 cm^{-1}) and phenylalanine (1004 and 1032 cm^{-1}). For the HWN region, the CH-stretching of lipids (2845, 2885 cm^{-1}) and proteins (2930 cm^{-1}) are the predominant molecular markers. Because of high cell density and the abundance of myelin lipids are the main constituents of human brain tissue.[215], [223] The $\text{CH}_2\text{-CH}_3$ deformation at 1441 cm^{-1} often dominates the spectrum. The signal from amide bands is also important, but strongly overlaps with lipid bands.[46], [49], [172] Markers of glioma include the phenylalanine band at 1004 cm^{-1} , the nucleic acids/ $\text{CH}_2\text{-CH}_3$ /Amide III band at 1339 cm^{-1} and the carotenoid bands at 1159 and 1523 cm^{-1} . [46], [49], [117], [172], [177], [189], [192], [223], [224] For other bands, studies show conflicting associations. In some cases, nucleic acid signal is strongest in tumor and necrotic tissue, whilst others demonstrate a decrease in malignant regions. The disagreements on how Raman signal changes as a result of pathological states stem from the multiple differences in experimental design from one study to the other (e.g. which types of tissue are compared, spatial resolution of the systems, tissue processing before Raman acquisition, analytical methods used to process and compare the Raman signals), along with low sample sizes (often less than 10 patients).[225]

Peak ratios are also considered to effectively summarize the lipid-to-protein content of a sample, with good discriminating power between white and grey matter, and normal brain and cancer. In these cases, involved bands included the 1442:1662 cm^{-1} (or 1441:1659 cm^{-1}) peaks,[117], [226], [227] 1299:1439 cm^{-1} peaks,[46] the 1266:1300 cm^{-1} peaks[228] and the 2930:2845 cm^{-1} peaks.[69], [99] Although simple and easily interpretable, these methods of spectral interpretation are limited by their inability to model complex interactions between the molecular species present in the living tissues.

As presented in the previous section, Raman linear unmixing (estimating the relative contribution of different pure endmembers to the total Raman signature) can help recover the tissue's biological content. Using VCA or NFINDR algorithms, authors concluded that cholesterol esters, nucleic acids, collagen and carotene contributions were higher in high-grade tumor while general lipid content and both lipid-to-protein and lipid-to-DNA ratios were decreased in malignant tissue.[115]–[117], [219] However, the reliability and consistency of this approach in unveiling the relative quantity of the molecular compound in the interrogated sample has not been thoroughly

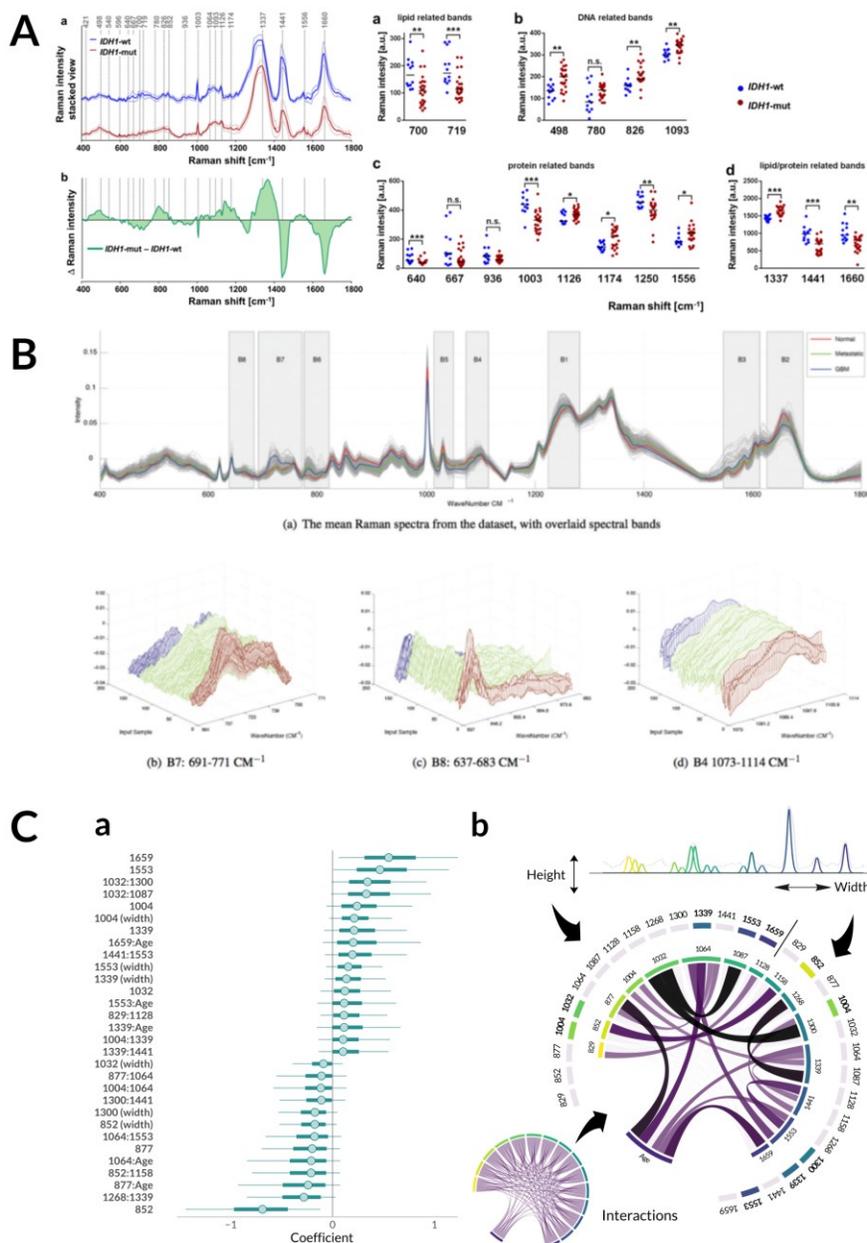


Figure 3.7: Chemometrics analysis of Raman data. **A:** Comparison of IDH mutation status in samples of human glioma. Left: the difference spectrum (green) between IDH-wt (blue) and IDH-mut (red) averaged spectrum is used to identify important Raman markers. Right: univariate analysis of the distribution of Raman intensity values at different Raman shifts. From Uckermann et al., 2018. Reproduced by permission from Springer: Nature. **B:** Spectral shape analysis of normal brain, glioma and metastatic human samples. The target regions are selected (a). For each region, parameters describing the shape of the distribution of Raman intensities

inside the regions are extracted and compared between the three classes (red: normal, green: metastatic, blue: glioma). From Stables et al., 2016. Reproduced by permission from the Royal Society of Chemistry. **C**: Analysis of importance of spectral markers between high-density and low-density/normal brain samples from *in vivo* human brain tissue. (a) the coefficients of a multivariate linear model are plotted. Features with a positive coefficient value are more prominent in samples with high-cancer density, while negative values are associated with low or absent cancer density. (b) Visual representation of the spectral markers (peak height, peak width, peak-to-peak and peak-to-patient-age interactions). From Lemoine et al., 2019.

Validated. It does rely on two important assumptions that may not hold true in biomedical applications: 1) the dataset of Raman spectra contains signal from pure molecular endmembers, and 2) the quantity of each molecular endmembers is independent of the others.[217], [218]

In reaction to this limitation, new approaches to the interpretation of biological Raman data have emerged that embrace the complex structure of the data-generating process and integrate into their model the uncertainties around the Raman signal and sample diagnosis. In Stables *et al.*, the authors selected spectral regions with high discriminating yield between glioma and normal brain samples, for which they extracted markers of Raman intensity distribution (**Figure 3.7B**).[188] These features carry more biochemical significance and stability than signal at a single Raman shift. The authors also calculated the ratio between every pair of extracted features as an additional feature ensemble, expanding the idea of interactivity between spectral information to all available variables. Their results demonstrated that the 691–771 cm^{-1} spectral region associated with phospholipids and amino acids had the highest discriminatory power, followed closely by the 637–683 cm^{-1} (nucleic acids and amino acids) and 1073–1114 cm^{-1} (nucleic acids and phospholipids). However, the samples were formalin fixed paraffin processed (FFPP) and Raman signal was altered to remove the paraffin-associated bands, which could explain some discrepancies between their findings and findings from fresh or frozen tissue. Nevertheless, this work was the first to incorporate systematic interactions between spectral feature as a potential discriminator between tissue classes, a key step in bridging the gap between classical Raman analysis and more complex computational pipelines used in other data-driven technologies such as neuroscience, genomics and proteomics.

In Lemoine *et al.* (Leblond's group), a feature engineering process integrating domain-specific knowledge, band-fitting and Bayesian optimization was applied to a dataset of *in vivo* human spectra to identify key features that differentiated normal or low-cancer density tissue from dense glioma.[195] These features included nucleic acids and protein bands, mainly collagen, phenylalanine and tryptophan. Furthermore, features that were generated by the authors revealed important discriminating power: pairwise band interactions, and interactions between patients' age and nucleic acid bands. Importantly, they were able to quantify the uncertainty around the effect size of each feature; this is especially valuable as chemometric analysis have shown many discrepancies between different groups.

Other examples of complex analytical techniques include graph-network representations as a tool to gain insight into the data's hidden structure, Bayesian statistics to model uncertainty in data acquisition and interpretation, and deep neural architecture to capture the complex hierarchy of the data-generating processes. Applied to Raman spectroscopy, these emerging techniques will likely expand the yield and depth of this technology and offer richer information to researchers and clinicians.

3.1.1.8 Future prospects in neurosurgery and neuroscience

3.1.1.8.1 Registering optical information in neuronavigation systems

In neurosurgery, optical measurement registration will be a critical aspect in the successful integration of Raman into the treatment protocol. Neuronavigation with infrared trackers and registration to preoperative MRI has been widely adopted by neuro-oncological surgeons. Registration of the Raman measurement to MRI images is an important aspect of future developments in intraoperative vibrational spectroscopy. The potential combination of MRI markers (such as distance to contrast-enhancement, T2- or T1-weighted intensity, apparent diffusion coefficient) with Raman-based markers could enhance the navigating environment for the operating team. In addition, MRI markers could increase the clinically relevant information content during the labeling phase of the Raman experiments. Spectroscopic measurements are meant to complement a panoply of factors that influence the extent of resection and minimize damage to normal brain; therefore, they should seamlessly integrate with the other available modalities.

In functional neurosurgery the challenge is different, and while similar to biopsy guidance, more difficult due to the small target areas. In DBS, for example, as the probe descends along a planned trajectory, the position of the optical measurement must be used to update the pre-operative imaging to account for any head-frame movement or brain shift.[229] Both of these steps would require electronic drive tools to perform the physical descent in order to automatically correlate probe depth with the optical measurement. While these electronic drive systems do exist for DBS surgery, manual drive screws are often used, in lieu, as they are faster. Since the information would ideally be merged into the current technologies that surgeons use, partnerships with commercial providers for stereotactic planning (*i.e.* Medtronic, Boston Scientific, Abbot) would greatly facilitate implementation in the operating room.

3.1.1.8.2 Outlook: Neurosurgery

At the present point in time, we may be nearing a clinical revolution in which pathology no longer requires visual confirmation by a trained clinician on site. We are seeing examples of this in the ocular industry where AI systems for diagnosing ocular pathologies are obtaining FDA approval.[230] While these systems still remain tools for the physician to aid in locating abnormalities, this may not always remain the case.

In the case of neuropathology, the diagnosis could soon be achieved intraoperatively using optical techniques, greatly decreasing traditional diagnostic turnaround time. In such a situation, optical modalities will compete for a share of the biomedical market and will aim to provide the most advantage to the surgeon. Raman's key advantage in this respect is the amount of information it can provide. The fingerprint type spectra can be used to quantify molecular ratios and discriminate tissue types in both a single spectrum format and in the form of content-rich hyperspectral images. Moreover, these capabilities are only beginning to be fully realized in intact brain tissue. As measurements are accumulated and sophisticated data science systems evolve for this application, a whole new Ramanomics field could emerge.[231] To achieve this, data sharing will become imperative. Properly labelled raw Raman spectra databases are extremely scarce in comparison to other fields and this is something that must change if the applied data science is to advance at a similar rate as other "omic" fields.

Whilst H&E stained slices are employed as "gold-standard" for labeling of Raman data, most recent guidelines on primary brain cancer classification are based on various other tissue markers

such as IDH mutation status, which has proven a unique and critical factor in establishing a prognosis orienting treatment for glioma patients[232]. Going beyond H&E staining and understanding how Raman signal changes as a function of these new biomarkers will be an important challenge of future Raman studies that aim to translate vibrational spectroscopy as a clinically valid decision-making tool.

The display of information is also an integral discussion point. In the case of Raman spectra, the raw optical data are much less important to the surgeon than what the optical information means. In the case of tissue discrimination using a point probe, a simple formulated label may suffice along with a metric for certainty. In the case of histology, SRH is already capable of displaying information in a way that would be familiar to the pathologist.[101] Whilst the information may evolve to provide more than classical H&E staining, this step is absolutely mandatory for clinicians to fully understand and accept the new technology.[233] In the case of functional surgery guidance, the live calculated position of the electrode overlaid on the pre-operative MRI would likely be the ultimate goal.

3.1.1.8.3 Outlook: Neuroscience

The information that is acquired with Raman spectroscopy is complex and difficult to interpret; therefore there is still much to be revealed in terms of through fundamental and clinical neuroscience research. There are a number of early stage studies showing the capacity of Raman to image or sense disease biomarkers such as prion proteins, amyloid beta plaques (Alzheimer's), alpha synuclein (PD), and even neurotransmitters which can be deficient in many neurological and psychiatric disorders.[98], [234], [235] This work has yet to be extended to human brain tissue *in vivo*.

In the future of Raman guidance for functional neurosurgeries, we could imagine the ability to measure the relative quantities of biomarkers and therefore the stage of the disease, to help guide treatment parameters. For instance in Parkinson disease, the loss of dopaminergic neurons in the substantia nigra results in a decrease in neuromelanin.[236] If Raman is used to guide DBS surgery in the future, it could also be used to measure either dopamine or neuromelanin concentrations to give information about the type and stage of the disease.[237]

Aging—more specifically brain-age—is another interesting topic for Raman spectroscopy in neurosurgery. Considerable research has gone into using MRI scans to show the relation between

structural changes and aging thanks to its non-invasive large-volume imaging.[238], [239] From this standpoint, Raman could be useful to complement and understand these observed trends, especially in the case of region-specific lipid changes associated with brain aging.[240] Furthermore, as there is likely an age-dependent change in Raman signatures, this type of data will be critical in age-matching Raman measurements to make them even more accurate for discrimination.

Looking to the far future, if the ever-growing field of optogenetics is ever applied to humans, this would imply the implantation of chronic fiber optics within a patient. If this were to come to fruition, maintained acquisition of Raman measurements from within the brain during stimulation could be possible. Although this would either require considerable downsizing of equipment or an optical fiber port interface on the skin, this remains an enticing prospect for the future.

3.1.1.9 Conclusion

Raman spectroscopy can provide label-free biomolecular information rapidly in a non-invasive manner and has the potential to revolutionize both neurosurgery and neurological research. The laser-based nature of Raman allows it to be incorporated into point probes, biopsy needles and microscopes, enabling its integration into multiple points of the neurosurgical workflow. Specifically, RS has the potential to improve treatment outcomes by aiding in the detection and delineation of healthy and cancerous tissues, blood vessels, and perhaps even disease-specific biomarkers. While this is promising, designing a Raman system for neurosurgical applications demands significant technical considerations both in terms of hardware implementation and data science methods, as is summarized in this article. It is our hope that this resource helps guide future developments in RS systems for neurosurgery.

3.1.1.10 Conflicts of interest

K. P. and F. L., are co-founders of ODS Medical Inc, a medical device company that seeks to commercialize the Raman spectroscopy system for real time detection of tissue abnormalities.

3.1.1.11 Acknowledgements

This work is supported by the Discovery Grant program from Natural Sciences and Engineering Research Council of Canada (NSERC), the Collaborative Health Research Program (CIHR and

NSERC), the Canada First Research Excellence Fund through the TransMedTech Institute and MITACS.

3.1.1.12 Tables

Table 3.1: Optical characteristics of neurosurgical point-probes. Abbreviations: Avg.: Average, Res.: resolution. Notes: *Number of wavelengths: 10, corresponding to ref.[66] **Average of 3s acquisition time.[45], [69] Irradiance (I) is calculated using the formula: $I = P_{avg} \times A$, where P_{avg} is average power, and A is the area of the beam spot on the tissue. The higher limit of the average power range is used for each case. Radiant exposure (H) is calculated using the formula: $H = I \times t$, I is irradiance and t is the acquisition time (ie. the time of light exposure.)

References	Specimen	In vivo	Modality	Raman Excitation Wavelength (nm)	Raman Shift (cm ⁻¹)	Excitation Spot Size (mm)	Avg. Optical Power (mW)	Irradiation (W/cm ²)	Acquisition time (s)	Radiant Exposure Single Acquisition (J/cm ²)	Spatial Res. (mm)	Spectral Res. (cm ⁻¹)
2007[104]	Porcine	No	SR	719	2400–3800	0.3	80	113	10	1131	~1	8.0
2010[46]	Mouse	Yes	SR	830	900–1800	0.4	120	95	5	477	~1	4.0
2015[43]	Human	Yes	SR	785	381–1653	0.5	37–64	33	0.15	4.9	~1	1.8
2017[44]	Human	Yes	SR, (AF, DRS)	785	381–1653	0.5	27–75	38	0.15	5.7	~1	1.8
2018[66]	Macaque	No	CR	792, 1030–1044	2800–3050	0.01	80	1.0×10^5	$0.001 \times \lambda$	1019*	< 0.01	1.0
2018[69]	Human	Yes	SR	671	2800–3600	0.35	10	10	2.0–6.0	31**	~1	1.8
2019[45]	Human	Yes	SR	671, 785	400–1800, 2800–3600	0.35	20	21	1.1–6.3	62**	~1	1.8

Table 3.2: System characteristics of neurosurgical point-probes. Abbreviations: OD: Outer diameter, SM: singlemode, MM: multimode, N/A: Not available.

Year	Probe Encasement	Probe OD (mm)	Autoclave Ready	Optical Setup	Detector	Optical Transport	Probe Description	Purpose
2007[104]	None	0.33	Yes	Uncased probe Free space optics	-70°C Cooled CCD (Renishaw)	Silica	1×300 μm MM source Same Fiber collection No filters	Research probe for brain tissue classification in neurosurgical guidance
2010[46]	Details N/A (Sedi-ATI)	1.6	Yes	All fibered System on cart	Details N/A, CCD (Horiba)	Silica	1×400 μm MM source 9×200 μm MM collectors No filters Lower AF due to 830 excitation	Stereotactic probe for tumor delineation in mouse cancer models
2015[43]	Stainless steel needle tybe (EmVision LLC)	2	Yes	All fibered System on cart	-40°C Cooled CCD (Andor)	Silica	1×200 μm MM source 7×300 μm MM collectors Micro-optic emission filter Micro-optic collection filter Custom lens	Handheld probe for tumor delineation and guided resection
2017[44]	Stainless steel needle tube (EmVision LLC)	2	Yes	All fibered System on cart	-40°C Cooled CCD (Andor)	Silica	1×272 μm MM Raman source 7×300 μm MM collectors 1×300 μm MM AF/DRS source 1×300 μm MM AF/DRS collector Micro-optic emission filter Micro-optic collection filter Custom lens	Multimodal handheld probe for tumor delineation and guided resection for improved cancer classification
2018[66]	Stainless steel needle tube (In-house)	0.4	Yes	All fibered System on cart	Photon Counter (H8259-02, Hamamatsu)	Silica	1×125 μm SM source 1×125 μm MM collector Optional GRIN lens for distal focusing	Modified DBS stylet for electrode implantation trajectory measurements in macaque cortex
2018[69]	Modified biopsy canula (Medtronic, Inc)	2	Yes	All fibered System on cart	-40°C Cooled CCD (Andor)	Silica	1×125 μm MM source/collector Angle polish for side viewing No micro-optical components required	Modified biopsy needle canula housing fiber-optic probe for biopsy guidance
2019[45]	Modified biopsy mandarin (EmVision LLC)	0.9	Yes	All fibered System on cart	-80°C Cooled CCD (Andor)	Silica	1×100 μm MM source 12×100 μm MM collectors Micro-optic emission filter Micro-optic collection filter Prism for side-viewing	Modified biopsy needle mandarin housing fiber-optic probe for biopsy guidance

Table 3.3: Optical Characteristics of state-of-the-art imaging systems for neurosurgery. Notes and abbreviations: N/A: Not applicable or not available, * : Hard to distinguish from publication, estimates are used. Peak irradiance and single pixel radiant exposure are calculated from the pulse peak power and the airy disk spot size, using the N.A. and wavelength of the published system (Table 2b) to derive the spot size from the Airy disk formula ($r_{\text{airy}} = \frac{1.22 \times \lambda}{2 \times NA_{\text{obj}}}$). Image area irradiance and radiant exposure use the same equations as in table 1a. For the radiant exposure calculation, total image acquisition time is use from table 2b. The wavelength used for all calculations was the one with the highest average power. For SRS this is the higher wavelength and for CARS this is the lower wavelength. It is critical to note that these values are extremely variable with respect to the optical arrangement, however, they can be useful to consider for clinical systems which aim to image intact tissue.

Purpose	Reference	Excitation Wavelengths [Pump, Stokes] (nm)	Power Sample Surface (mW)	Rep Rate (MHz)	Pulse Width (ps)	Pulse energy Single Pulse (nJ)	Peak Power Single Pulse (W)	Peak Irradiance Single Pulse (W/cm ²)	Pixel Dwell Time (μs)	Irradiance Image Area (W/cm ²)	Radiant Exposure Image Area (J/cm ²)
Rapid Histology	2017[101], [129]	[790, 1010–1040]	[120, 140]	40	2	6.5	3250	2.62×10^{11}	2	163	327
	2018[131]	[680–1300, 1041]	[20, 40]	80	2	0.75	375	3.0×10^{10}	2.4	234	6900
	2019[123], [241]	[781, 1005]	[50]*	80*	1.2	0.625	520	7.3×10^{10}	0.4	220	4.5
Handheld Imager	2018[145]	680–1300, 1041]	[40, 40]	80	2	1	500	9.9×10^9	6	283	850
	2018[132]	785	500	N/A	N/A	N/A	N/A	N/A	N/A	3.5	300
Endoscope	2017[140]	[816, 1064]	[40, 40]	80	6	1	167	5.7×10^7	N/A	88.9	890
	2018[141]	[800, 1040]	[20, 10]	80	0.1	0.375	3750	1.0×10^{15}	N/A	29.3	23.4
	2018[142]	[817, 1064]	[60, 35]	76	6	1.25	208	6.9×10^9	6	989	376

Table 3.4: Physical Characteristics of state-of-the-art imaging systems for neurosurgery. Notes and Abbreviations: *Hard to distinguish from article, WD: Working Distance, OD: Outer Diameter, N/A: Not applicable or Not available. Total image acquisition time was calculated using single band imaging time multiplied by the number of bands imaged.

Purpose	Year	Specimen	Modality	Portability	Orientation	Reported Spatial Res. [Lateral, Axial] (μm)	Objective NA	FOV (μm)	Wavenumber Acq. Time (s)	Spectral Range (cm^{-1})	Spectral Res. (cm^{-1})	Total Acq. Time (s)
Rapid Histology	2017[101], [129]	Human Fresh resection	SRS	Fibered laser On cart	Transmission	[0.36, 1.8]	1	400×400	2	2845, 2930	N/A	4
	2018[131]	N/A* Snap frozen	SRS	Free space Optic table	Backward	[0.45*, N/A]	1	160×160	0.95	2800–3000	15	3
	2019[123], [241]	Human Fresh	CARS, TPEF, SHG	Free space Optic table	Backward	[0.5, N/A]	1	150×150	0.02	2845	N/A	0.02
Handheld Imager	2018[145]	Canine Snap frozen	SRS	Handheld Optic table	Backward	[1.4, N/A]	0.5	168×168	0.125	2800–3000	15	3
	2018[132]	Calf Fresh	SR	Handheld (2cm WD) On cart	Backward	[100, N/A]	N/A	4000×3500	0.7	940–1800	120	84
Endoscope	2017[140]	Skin	CARS, TPEF, SHG	8mm OD Endoscope Optic table	Backward	[10, N/A]	0.5	$\pi \times 150^2$	10	2845	N/A	10
	2018[141]	Colon	CARS, SHG	4mm OD Endoscope Optic table	Backward	[0.8, 5.9]	0.45	320×320	0.8	2885	N/A	0.8
	2018[142]	Human nerve Cryosection	CARS, TPEF	2.2mm OD Endoscope Optic table	Backward	[1.4, N/A]	0.5	310×310	3.8	2841	N/A	3.8

Table 3.5: Feature engineering and supervised machine learning algorithms for classification of Raman spectroscopy in brain tissue

Task	Technique	Description	Pros	Cons
Feature Engineering	Principal Component Analysis	Unsupervised algorithm. Extracts the projection of the data matrix onto a set of linearly uncorrelated Principal components	Agnostic to tissue classes (unbiased) Effective feature reduction Fast Removes correlated (redundant) features Effective visualization technique	Expert intervention is impossible Agnostic to tissue classes (can reduce classification performances) Susceptible to features unrelated to the classification task (e.g. noise and artifacts)
	Feature extraction	Generate new variables from the spectra (e.g. band-fitting) or by transformation of already available variables (e.g. peak ratios)	Increases the information content of the data Interactions are accounted for Better captures molecular processes	Requires expert knowledge Time and effort-intensive Requires high data quality
	Feature selection	Identification of the variables which show highest correlation with class (filter method). Can be performed as a second feature-engineering step, after PCA or feature extraction	Simple Task specific	Biased Prone to false positives
Supervised Learning	Linear discriminant analysis	Estimates a set of multivariate normal distributions that better explain the data and assigns new observations to the class with highest likelihood	Simple Fast Low data requirements	Assumes normality of distribution Assumes homoscedasticity Requires independent variables
	Support Vector Machine	Finds a hyperplane that maximizes the margin between the support vectors, i.e. the observations of each class closest to the hyperplane	Adaptable to non-linear feature spaces Fast when N is small Easily implemented	Performance is highly dependent on hyperparameters Requires engineered features
	Decision trees	Non-parametric model that assigns a decision function to each variable. Decision trees can be aggregated to increase stability by averaging (Boosted trees) or training over many subsamples (Random Forests)	Intuitive Able to model complex feature spaces	Unstable by itself Aggregation is time consuming
	Artificial Neural Networks	Connectionist model that considers each variable as an input neuron. These neurons are subsequently connected to a pre-defined number of hidden layers through an activation function. These hidden layers connect to an output layer that generates the class prediction. ANN with more than one hidden layer are called Deep Neural Networks	Adapts to any non-linear function Highly customizable High performances Does not require feature engineering (Deep Neural Networks)	High data requirements Performance is highly dependent on hyperparameters Computationally intensive

3.2 Statistical Learning

Cutting-edge Raman spectroscopy systems can capture the rich molecular signatures of *in vivo* tissue samples in real-time. The resulting spectra are high-dimensional, and their features are highly interactive. These attributes greatly complicate the analysis of spectroscopic data; yet, they could be addressed by leveraging simple computational techniques. In this section, key concepts in statistical learning are reviewed based on the most up-to-date literature on the subject. This chapter serves to give a foundation in statistical inference and Bayesian statistics, and identify analytical strategies that could benefit Raman spectroscopy of *in vivo* brain tissue.

3.2.1 Statistical models

In attempt to use experimental data to answer a scientific question about a population of interest, researchers are faced with two important restrictions: they cannot access all information on every individual, and they cannot study every individual of the population. This has two consequences:

- 1) Researchers need to define a *mathematical model* that will represent the data, or the reality into which our data is defined. Models allow to simplify our representation of the population and make the studied system amenable to mathematical operations. A model can be very simple, such as a statistical distribution with one or two parameters, or highly complex like a hierarchical network with thousands of parameters.
- 2) The results of a statistical analysis (for example, the parameters of the statistical distribution we selected as model for our data) can only apply to the sample that was under observation. *Statistical inference* is the generalization of this result to the entire population of interest, beyond the observed sample. Inference results in an *estimate* of the real-world. Estimations comes with a certain probability of error, which is also estimated by the researcher.

Experimental data consists in collected information on several individuals. Each distinct individual is called an *observation* or an *example*; the number of observations is represented by the variable n . In Raman spectroscopy, a single Raman acquisition is an observation. For each observation, specific attributes are collected, referred to as *variables*, *predictors* or *features*. The number of

variables is represented by p . For Raman spectra, the intensity values at each Raman shift are the variables. The data are stored in a matrix $\mathbf{X}_{n \times p}$.

Supervised statistical learning consists in predicting an outcome—or dependent variable—from a set of predictors—or independent variables. The outcome could be the nature of the interrogated tissue sample (*e.g.* “cancer” vs. “normal brain”), and the predictors, the intensities of the Raman spectrum issued from that sample. As explained in the previous chapter, assuming there exist a random variation ϵ that cannot be explained by the predictor variables, this type of problem has the form $\mathbf{y} = f(\mathbf{X}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}$, with \mathbf{y} the vector of outcomes, \mathbf{X} the matrix of variables, $\boldsymbol{\epsilon}$ a vector of random error terms for each observation and f the model describing the relationship between predictors. The parameter vector $\boldsymbol{\theta}$ will be estimated based on available data; this process is called the model’s *training*.

One of the most common and simple form of learning models is the linear model[198], [242]. A linear model assumes that y can be predicted by the additive combination of the vector of variables $\mathbf{x} = (x_1, x_2, \dots, x_p)$. It has the form:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \epsilon. \quad (3.1)$$

The intercept α and coefficients $\boldsymbol{\beta}$ are the parameters of the model. Given a dataset \mathbf{X} and outcome vector \mathbf{y} , we can find the values of α and $\boldsymbol{\beta}$ that minimizes the error vector $\boldsymbol{\epsilon}$. We can re-write the relationship between error term and outcome as $\mathbf{y} = \hat{\mathbf{y}} + \boldsymbol{\epsilon}$, where $\hat{\mathbf{y}}$ is the estimated value of the outcome given the predictors. **Figure 3.8** illustrates the difference between each estimated value \hat{y} (or predicted values) and the actual, observed value y of the outcome in a linear model with only one predictor variable x .

3.2.1.1 Link functions

When training a statistical model, the optimal value for α and $\boldsymbol{\beta}$ given the collected data are calculated by minimizing a loss function $L(\mathbf{y}, \hat{\mathbf{y}})$. Commonly, for linear models, the loss function is the mean squared error (MSE)[243]. MSE is defined by the formula:

$$MSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (3.2)$$

With MSE, the values of predicted outcomes \hat{y} are estimations of the **mean** value that the outcome could take given the predictors x , and the error terms ϵ the **standard deviation** of these estimated values[198]. In Raman spectroscopy, \hat{y} could be an estimation of the number of cancer cells in a sample, and x the sample's Raman acquisition. MSE is therefore convenient for most use cases, as the normal distribution of error terms around the estimated values \hat{y} is a reasonable assumption given the Central Limit Theorem. This theorem states that when random observations are added, their sum tends to converge to a normal (or gaussian) distribution—even though their individual distribution might not be normal[243].

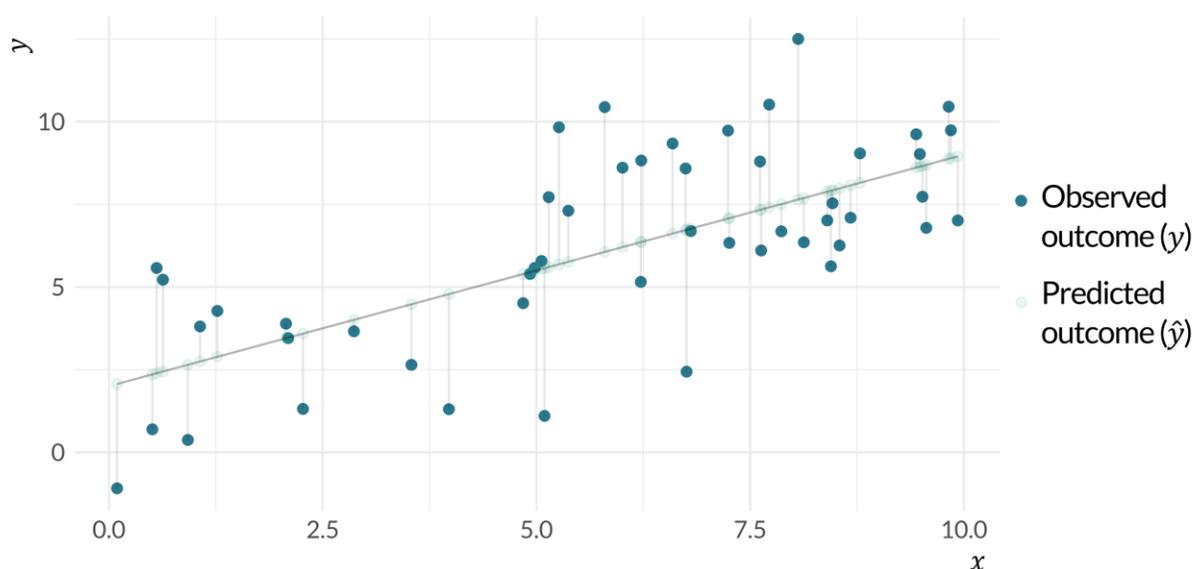


Figure 3.8: Linear model for a dataset with a single variable x and an outcome y . The model has the form $\hat{y} = \alpha + \beta x$. The error between the predicted and actual value of y is represented by the vertical lines.

The downside of MSE, however, is that the linear model is, in theory, only applicable to continuous outcome with normally distributed error. For example, if the outcome is binary, it can only assume two values (*e.g.* 0 or 1, success or failure, cancer or normal tissue). Simple linear regressions are unbounded; the values of predicted outcome \hat{y} can assume any values in the interval $[-\infty, +\infty]$. Linear models can be adapted to other probability distributions by adding a link function. In the case of binary data, the continuous outcome \hat{y} can be coerced into the $[0, 1]$ interval with the logistic function:

$$p = \frac{1}{1 + e^{-\hat{y}}}, \quad (3.3)$$

where p is the probability that the outcome has a value of one. The difference between a linear regression and linear regression with logistic link (Logistic regression) for binary outcome is illustrated in **Figure 3.9**.

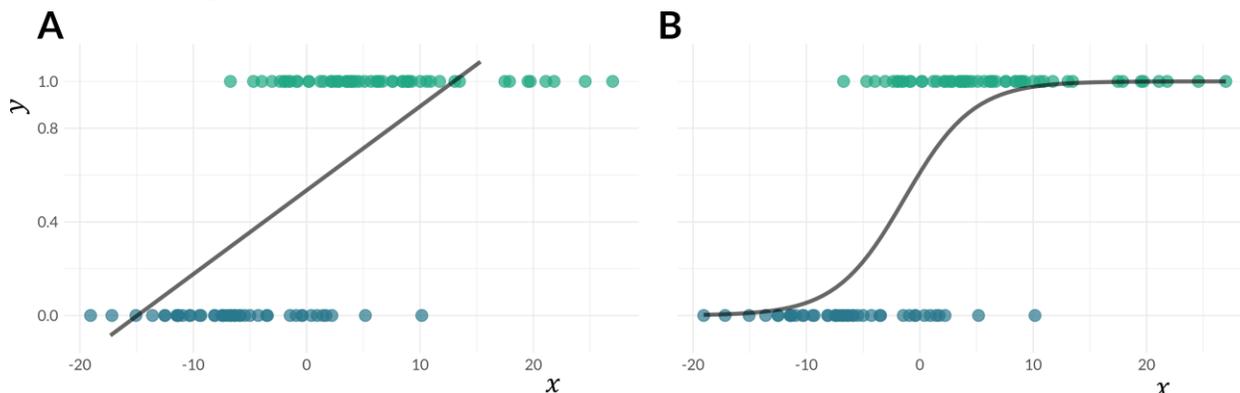


Figure 3.9: Modeling binary outcome with one predictor variable. A: While the outcome is binary (0 or 1), the linear regression line fits poorly to the data: for extreme values of x , the predicted \hat{y} will be above 1 or below 0. B: with a logistic link function, the predicted outcome is restricted to the range $[0,1]$.

3.2.1.2 Interactions

Linear relationships are simple to define and interpret. However, they are unlikely in the setting of a real-world problem. One of the reasons is that linear models assume that predictors are independent of each other, *i.e.* that changing the value of one predictor does not affect the effect of the others. In the real-world, the variables that shape a population are often intertwined. In Raman spectroscopy, a change in the molecular composition of the interrogated sample affects multiple Raman bands. Conversely, a single Raman band is often associated with a multitude of distinct molecular compounds.

This constraint of independent additivity between predictors can be relaxed by adding *interaction terms* to the model. In a model with two predictors x_1 and x_2 which are thought to interact together and an error variable ϵ , the linear equation with interaction can be defined as:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 (x_1 \times x_2) + \epsilon. \quad (3.4)$$

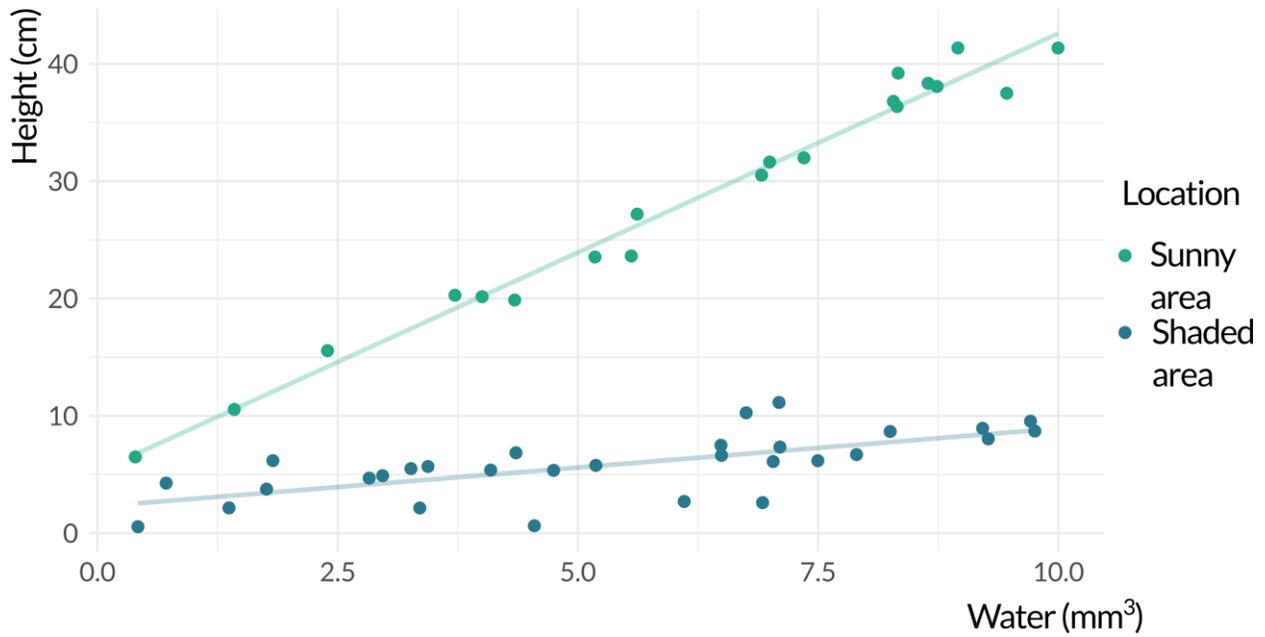


Figure 3.10: Simulated example to illustrate the effect of an interaction term in a linear model.

In this case, β_3 is the coefficient of the interaction terms ($x_1 \times x_2$). To better understand the meaning of the interaction, we can rearrange the terms in eq. (3.4) as:

$$y = \alpha + (\beta_1 + \beta_3 x_2) \times x_1 + \beta_2 x_2 + \epsilon. \quad (3.5)$$

Without the interaction term, an increase of x_1 by one-unit results in a constant change in y equivalent to β_1 , regardless of the value of x_2 . With the interaction term, however, the effect of x_1 on y is no longer constant: it depends on x_2 by a factor equal to β_3 .

We can simulate a hypothetical problem that will help illustrate the effect of adding an interaction term. In this simulation, we want to find the relationship between the height of plants at 6 weeks (in cm) and two predictor variables: the amount of water that the plant has received (in mm^3), and the location of the plant, either in a sunny area (1) or shaded area (0). We collect the data for $n = 50$ plants in the data matrix $\mathbf{X}_{50 \times 2}$ and the height of the plants in the outcome vector $\mathbf{y} = (y_{\text{plant } 1}, y_{\text{plant } 2}, \dots, y_{\text{plant } n})$. We could pose the problem as a linear regression model with two independent variables. Instead, we hypothesize that plants in a shaded area will not benefit from an increased amount of water as much as plants in a sunny area. Therefore, we include an interaction term into our model:

$$y = \alpha + \beta_{\text{water}} x_{\text{water}} + \beta_{\text{sun}} x_{\text{sun}} + \beta_{\text{water:sun}} (x_{\text{water}} \times x_{\text{sun}}) + \epsilon. \quad (3.6)$$

The result of this simulated experiment is illustrated in **Figure 3.10**. The *slope* of the regression line of y on x_{water} varies depending on the value of x_{sun} . In this simulation, $\beta_{water:sun} = 3$. Therefore, according to eq. (3.6), the effect of an increase in 1 mm^3 of water on the height of the plant is 3 times more important when the plant grows in the sun than when it grows in the shade.

3.2.2 Overfitting, underfitting and regularization

Linear models are popular because they are simple. They have great interpretability, because the effect of each variable can be calculated via the value of its associated coefficient. And while they tend to be inaccurate and oversimplified representation of natural phenomenon, they tend to preserve a good generalizability outside of the studied sample[198].

Complex models will closely fit the data on which they are trained. The training sample could only represent a specific subset of the studied population, either because of systematic biases during the data collection, or even by chance. In this case, the performance of our model on the training data may not be a good representation of the performance on the model in the real-world. A more rigorous approach is to see how well a trained model fits an *out-of-sample* subset of the data, one with which the model's parameters were not estimated. This process is called *testing* the model

During model training, the loss function $L(\mathbf{y}_{train}, \hat{\mathbf{y}})$ is minimized, resulting in a function \hat{f} that is an estimation of the real-world data-generating function f . \hat{f} can then be used to generate new estimates $\hat{\mathbf{y}}$ given new examples $\mathbf{X}_{test} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ that are associated with outcome vector \mathbf{y}_{test} . The distance between \mathbf{y}_{test} and $\hat{\mathbf{y}}$ is the *testing error*, and it composed of three elements: the bias of \hat{f} , the variance of \hat{f} and the irreducible error of \mathbf{X}_{test} (or ϵ_{test})[198].

Bias refers to a systematic error resulting from using a model which is simpler than the real data-generating process. For example, defining the relationship between the height of a projectile and the longitudinal distance it has traveled with the linear model $height = \alpha + \beta \times distance$ will undoubtedly result in large error, because the model is simply not appropriate for the underlying phenomenon.

On the other hand, **variance** is increased when a model is highly susceptible to very small changes in the training data. This can be the case of very high-order polynomial; a small change in the value of one of the observations can drastically affect the shape of the curve. Therefore, the fit of

the model is very dependent on the sample it is trained on. Given a very large sample that closely resemble the true population, variance is minimized and the function \hat{f} can be very close to f .

Finally, the **irreducible error** ϵ is only dependent on the data and is the minimal error that will remain given the perfectly estimated function f . This error is either caused by stochastic processes that cannot be estimated, or by deterministic phenomenon that were not identified or recorded during the data collection (*e.g.* an unobserved predicting variable).

Therefore, our choice of model will depend on the balance between bias and variance. Sometimes, we can afford a very flexible model because we have a very large sample of great quality; we are confident that there exists very little difference between our sample and the real-world population. On the other end, with small datasets or large number of variables, the model's variance becomes a concern. In biomedical Raman spectroscopy experiments, we have access to a limited number of tissue samples (small n), and for each sample, a Raman acquisition contains hundreds of variables (large p). In those cases, in order to improve the predicting performances of our model on external data, we can purposefully restrict its flexibility. This is called *regularization*. In applied statistical learning, regularization virtually always lead to a model with higher predictive performance[187], [244], [245]; we tend to underestimate the variance of the models we build.

Regularization in linear models can be performed by adding a *constraint* to the optimization problem. To estimate the α and $\boldsymbol{\beta}$ parameters of the linear model, we must minimize the MSE and solve:

$$\arg \min_{\boldsymbol{\beta}} \sum_{i=1}^N \left(y_i - \alpha - \sum_{j=1}^p x_{ij} \beta_j \right)^2. \quad (3.7)$$

Adding a constraint is equivalent to allocating a “budget” for the sum of the values of the coefficients $\boldsymbol{\beta}$. We define a regularizer $R(\boldsymbol{\beta})$ that will be added to eq. (3.7). There are two common constraints for linear models: L_1 and L_2 penalizations. In L_1 -regularization, $R(\boldsymbol{\beta}) = \lambda \sum_{j=1}^p \beta_j$; for L_2 -regularization, $R(\boldsymbol{\beta}) = \lambda \sum_{j=1}^p \beta_j^2$ [246]. In both cases, the parameter λ is a scaling factor that determines the strength of the regularization[187].

A regularizer will reduce the value of the model's coefficient. The L_1 penalty, also called the *Lasso*, has an interesting property: it will result in some of the coefficients being exactly equal to

zero[187]. Depending on the value of λ , only a subset of variables will influence the predicted outcome. This property becomes useful when we need to perform feature selection, *i.e.* identify which features (*e.g.* which parts of the Raman spectrum) are the most important to predict an outcome. In the Lasso, the optimization problem becomes:

$$\arg \min_{\beta} \left\{ \sum_{i=1}^N \left(y_i - \alpha - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p \beta_j \right\}. \quad (3.8)$$

3.2.3 Frequentist and Bayesian inference

Earlier, we described statistical inference as the generalization of an experiment to the entire population, beyond the studied sample. Multiple inference frameworks exist; among them, *Frequentist inference* is, by far, the most widely used. A frequentist analysis rests on the falsification of the null hypothesis, which states that there is an absence of effect in the population where the studied sample is drawn. To achieve this, a frequency statistic, the p -value, is calculated: if p is small enough, frequentists can reject the null hypothesis. The p -value is the fraction of times we would get an equal or larger effect than measured during our experiment if we were to repeat the experiment an infinite number of times. Before the statistical analysis is performed, a confidence level is arbitrarily defined (classically: 5%): this is the threshold frequency above which the null-hypothesis is still plausible. If the null-hypothesis is rejected, the frequentist concludes that there is indeed a “statistically significant” effect[247].

Recently, the frequentist’s p -value received a lot of negative attention from the scientific community. Statisticians view the p -value as reductionist, biased, and susceptible to misinterpretation[248]–[250]. P -values are point estimates based on an assumed probability distribution that depends only on the size of the effect and the size of the population; other factors such as prior evidence, study design and confidence in the acquired data are left out of the equation. Also, their capacity in rejecting the null-hypothesis depends heavily on the false-positive risk of the experiment (see [251]). More importantly, frequentist statistics are often misinterpreted in terms of probabilities[252]–[255]. For example: “A p -value of 0.01 means that there is 1% probability that the observed parameter is due to chance”, or “the 95% confidence interval is the range of values which has a probability of 95% of containing the true population parameter”. While

this interpretation is much more intuitive, it is incorrect; frequentist analysis results are based on deterministic, true parameters value, and do not allow for probabilities[247].

The real-world population, however, certainly contains stochastic processes; we can reasonably assume that there is no single “true” parameter value that can be exactly calculated. There is a level of uncertainty in everything we measure, and this concept is at the core of Bayesian analysis[256]–[258]. *Bayesian inference* aims to estimate the conditional probability of a parameter θ (*i.e.* the varying degree of likelihood of θ at different values), knowing the data. Bayes theorem is as follow:

$$P(\theta|\text{data}) = \frac{P(\text{data}|\theta) \times P(\theta)}{P(\text{data})}. \quad (3.9)$$

The term $P(\text{data})$ is a normalization constant that does not depend on $P(\theta)$. As we are not interested in the absolute likelihood of $P(\theta)$, but rather in its conditional probability at different values, eq. (3.9) can be simplified:

$$P(\theta|\text{data}) \propto P(\text{data}|\theta) \times P(\theta), \quad (3.10)$$

that is: the probability of the parameters given the data (*posterior*) is proportional to the probability of the data under different values of parameters (*likelihood*) multiplied by the probability of the parameters (*prior*). For a clear and famous example on how the frequentist and Bayesian approaches differ, see [259].

3.2.3.1 Likelihood, priors and posterior distributions

The main steps in a Bayesian analysis are: 1) defining a model for the data (the *likelihood*), 2) choosing *prior* distribution for the different parameters of the model, and 3) computing the *posterior* distribution of these parameters (**Figure 3.11**).

The **likelihood** is a function that describes the plausibility of the data. The choice of likelihood function depends on the assumptions made by the statistician about the data generating process. For example, in a coin tossing experiment where we would want to evaluate the possibility that the coin might be rigged, we toss the coin n number of times and count the number of times h where the toss results in “heads”. We could choose as likelihood function the binomial distribution:

$$P(h|n, p) = \frac{n!}{h!(n-h)!} p^h (1-p)^{n-h}, \quad (3.11)$$

where p is the parameter to be estimated. This model makes some assumptions: that every coin toss is independent of the other, and that each trial has the same probability of resulting in “heads”. By changing the value of p in the equation, we obtain different values for $P(h|n, p)$. For example, the likelihood that $p = 0.9$ in an experiment where the number of coin tosses $n = 10$ and the number of heads $h = 5$ is 0.0015 ($P(h = 5 | n = 10, p = 0.9) = 0.0015$). The value of p that maximizes $P(h|n, p)$ is called the *maximum likelihood estimation* (MLE) of p (in this example, the MLE of p is 0.5).

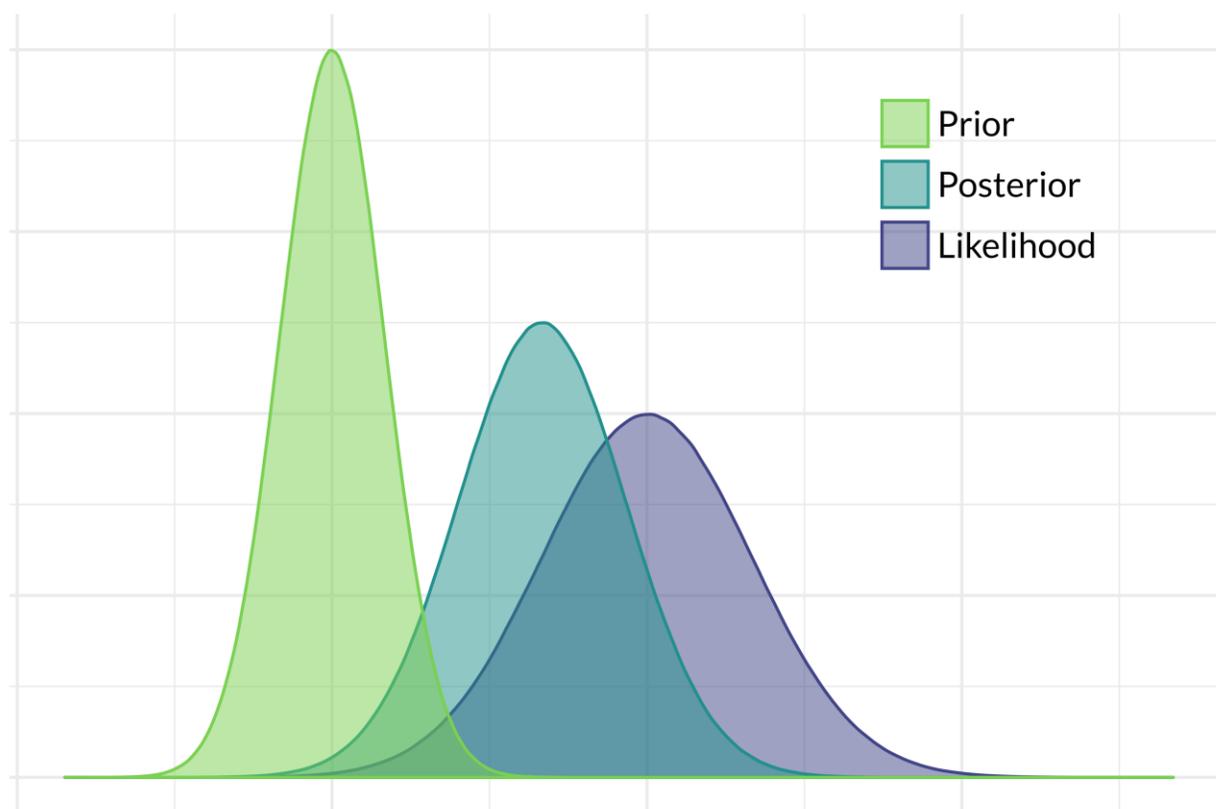


Figure 3.11: Depiction of the posterior distribution as the result of the prior distribution being updated by the likelihood, or collected data.

The **priors** are the initial set of possibilities for every parameter that the Bayesian algorithm will estimate. Priors allow to inject previous knowledge about the data into the analysis. For example, they could represent the results of previous experiments, or beliefs held by the statistician on the possible values that the parameters can take. Most of the time, however, priors are set to assist in the convergence of the algorithms and are agnostic to the data that is being analyzed[260], [261]. Prior distributions regularize the Bayesian analysis; just like constraint on optimization in

frequentist models, they restrict the search-space and avoid extreme values for parameters[262]. The choice of priors, as many other decisions in both Bayesian and Frequentist inference, depends on the researcher and the scientific problem, and must be rigorously justified[263].

Priors are described in terms of probability distribution. In the coin-tossing example above, a non-informative prior for the parameter p could be the uniform distribution between 0 and 1, *i.e.* $p \sim \text{unif}(0, 1)$. We could also start with the assumption that the coin is not rigged and define the prior as a normal distribution centered at $\mu = 0.5$: $p \sim \mathcal{N}(0.5, \sigma)$, where the value of σ is chosen by the statistician. Small σ values will make the posterior distribution tend towards 0.5, whereas large values will allow p to adopt a wider range of possible values. Therefore, the prior $p \sim \mathcal{N}(0.5, \sigma)$ acts as a regularization term, with the parameter σ controlling the strength of the regularization (analogously to the parameter λ in L_2 -regularization in: $R(\boldsymbol{\beta}) = \lambda \sum_{j=1}^p \beta_j^2$ [262]). Because μ and σ are related to the model and not the data, they are considered *hyperparameters* of the model.

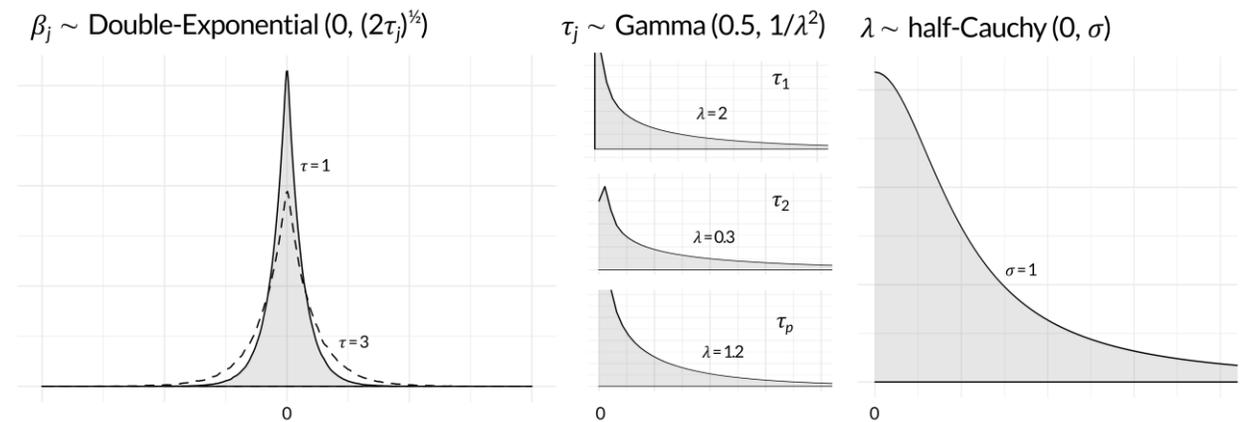


Figure 3.12: Bayesian Hyper-Lasso priors. The prior on the coefficient is a Double-Exponential. The parameter τ which regulates the tails of the distribution, varies for every coefficient in $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)$. The parameter τ is drawn from a Gamma distribution, for which the λ parameter varies for every coefficient and is drawn from a half-Cauchy distribution with $\sigma = 1$.

A basic regularizing prior for Bayesian models is the Double-exponential distribution, shown in the left panel of **Figure 3.12**. This distribution is heavily centered at zero; this has the effect of drawing the posterior distribution of the coefficients towards zero. Therefore, only the coefficient of variables with a very strong effect on the outcome will exhibit a positive or negative value. The distribution has a single parameter τ , that defines its tail. With Double-exponential priors, the MLE

of the coefficient will be equal to those that are obtained with L_1 -regularization (eq. [3.8])[264]; the approach was therefore named the “*Bayesian Lasso*”. Because the coefficient of uninteresting variables will tend towards zero, the Double-exponential priors can be used as a feature selection filter. Yet, in small samples and high-dimensional settings, the Bayesian Lasso is very restrictive and can fail to retrieve some variables with smaller effect size[262].

Recently, other approaches have been proposed to improve the Bayesian Lasso, ones that rely on a multilevel definition of the Double-exponential priors[265], [266]. In a multilevel model, hyperparameters have their own prior distributions. This can convey interesting mathematical properties to the priors and improve the convergence of the algorithms used to estimate the posterior distribution[267], [268]. Designing multilevel architectures that can regularize Bayesian’s analyses is a subject of active research[262]. Among multilevel priors, the Hyper-Lasso has shown the best performances in recovering important features in settings with a high number of variables and small number of observations, an important limitation of biomedical Raman experiments[269], [270]. The Hyper-Lasso sets a prior distribution on the τ parameter of the Double-Exponential (**Figure 3.12**). As such, τ varies for every coefficient, and important variables will more easily reach extreme values while unimportant variables will tend strongly towards zero. Each value in $\boldsymbol{\tau} = (\tau_1, \tau_2, \dots, \tau_p)$ is drawn from a Gamma distribution, for which the scale parameter θ is equal to $\frac{1}{\lambda^2}$. The value of λ is drawn from a half-Cauchy prior. The Hyper-lasso is described in the following equation:

$$\begin{aligned}\beta_j &\sim \text{Double-Exponential}\left(0, (2\tau_j)^{\frac{1}{2}}\right) \\ \tau_j &\sim \text{Gamma}\left(0.5, \frac{1}{\lambda^2}\right), \text{ for } j = 1, \dots, p \\ \lambda &\sim \text{half-Cauchy}(0, \sigma = 1).\end{aligned}\tag{3.12}$$

The **posterior distribution** is the updated joint probability distribution for all the model’s parameters. The posterior is the result of multiplying the prior distribution with the likelihood (or *updating* the priors with the evidence) and normalizing by the marginal probability of the data so that the function integrates to 1. This can be solved analytically only if the prior and likelihood are from the same family of probability distribution, imposing an important restriction on the model’s definition. With increasing computational capacities, it is now much more advantageous to

estimate the posterior distribution numerically[256]. Most algorithms that approximate posterior distributions are based on Markov Chain Monte Carlo (MCMC) sampling, introduced in the next section.

The following example shows how we would define a simple linear model in a Bayesian framework. Here, we want to estimate the *waist circumference* w in centimeters of patients in a hospital with one predicting variable, *age*. The Bayesian problem is defined as followed:

$$\begin{aligned}
 w_i &\sim \text{Normal}(\mu, \sigma) \\
 \mu &= \alpha + \beta_{\text{age}}x_{\text{age}} \\
 \alpha &\sim \text{Normal}(\mu = 102, \sigma = 40) \\
 \beta_{\text{age}} &\sim \text{Normal}(\mu = 0, \sigma = 10) \\
 \sigma &\sim \text{half-Cauchy}(x_0 = 0, \gamma = 10).
 \end{aligned} \tag{3.13}$$

The first two lines represent the likelihood. They are the model which puts in relation the data (x_{age}) with the parameters (α , β_{age} and σ). We assume that the waist circumference of the patients follows a normal distribution conditional on the age of the patients. The last three lines are the prior distributions for each parameter. The prior shown above represent weakly informative priors and will act to stabilize the sampling of the posterior distribution, avoiding the drawing of extreme values. They are based on previous knowledge (average waist circumference of the U.S. populations[271]). The resulting posterior distribution $P(\alpha, \beta_{\text{age}}, \sigma \mid x_{\text{age}})$ can be used for various applications:

- 1) Summarize the probability of each parameter. For example, describe the probability that β_{age} lies above a certain value, or estimate which value of σ has the highest probability.
- 2) Simulate and predict. Given the joint distribution for the parameters, we can provide new values for x_{age} to the model and output the corresponding vector $\hat{\mathbf{w}}$. The new x_{age} values can be obtained from real hospitalized patients, or they can be simulated to check the model's fit and behavior. Simulation and prediction can also be used to compare different models and choose which model is more suited to our specific application.

3.2.3.2 Markov Chain Monte Carlo

A Markov Chain is a sequence of random variable x for which the value of x^t (or x at time t) depends only on the value of x^{t-1} . In Bayesian inference, Markov Chain are used to guide the numerical approximation of the joint posterior distribution of unknown parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$. The samples are initially drawn from an initial distribution $P_0(\boldsymbol{\theta})$. At each iteration, this distribution is updated, so that it eventually converges to the true $P(\boldsymbol{\theta})$. The algorithms that achieve this convergence are usually derived from the Metropolis algorithm[272].

The Metropolis algorithm uses an acceptance/rejection decision at every step of the random walk[273]. An initial sample $\boldsymbol{\theta}^0$ is drawn from a starting distribution $P_0(\boldsymbol{\theta})$. Then, a candidate sample $\boldsymbol{\theta}^*$, drawn from a *jumping* distribution $J_t(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{t-1})$, is proposed. The proposed sample can be either accepted or rejected. The probability of acceptance depends on the relative likelihoods of $\boldsymbol{\theta}^0$ and $\boldsymbol{\theta}^*$ with respect to the model and data. If the proposal is accepted, $\boldsymbol{\theta}^*$ becomes $\boldsymbol{\theta}^t$. In the case of rejection, $\boldsymbol{\theta}^0$ becomes $\boldsymbol{\theta}^t$. The sequence of proposal continues until an acceptable number of samples has been drawn. The cumulation of all samples can then be plotted as a histogram to visualize the posterior distribution for each parameter in $\boldsymbol{\theta}$. More modern forms of MCMC offer very efficient proposals and require a smaller number of iterations before converging to the desired distribution[274], [275]. This is essential in the case of large, hierarchical models with hundreds of parameters.

3.2.3.3 Bayesian predictions

After having drawn the multivariate posterior distribution for the vector of parameters $\boldsymbol{\theta}$, we can use these values to perform predictions on new data. Following the Bayesian framework, predictions will not be point estimates, but entire predictive distributions $P(\hat{y}_i|\mathbf{x}_i, \boldsymbol{\theta})$ for every observation \mathbf{x} in the testing set (**Figure 3.13**). For a single observation \mathbf{x}_i , we can compare the real, observed value y_i against the estimated predictive distribution $P(\hat{y}_i)$, yielding the likelihood of y_i . All the individual likelihood for every observation $(\mathbf{x}_0, \dots, \mathbf{x}_n)$ of the testing set can be computed, and their product will result in the overall likelihood of the observed \mathbf{y} knowing the estimated $P(\hat{\mathbf{y}})$. This metric is used to assess the performance of a model and is detailed in Chapter 4.1.3.

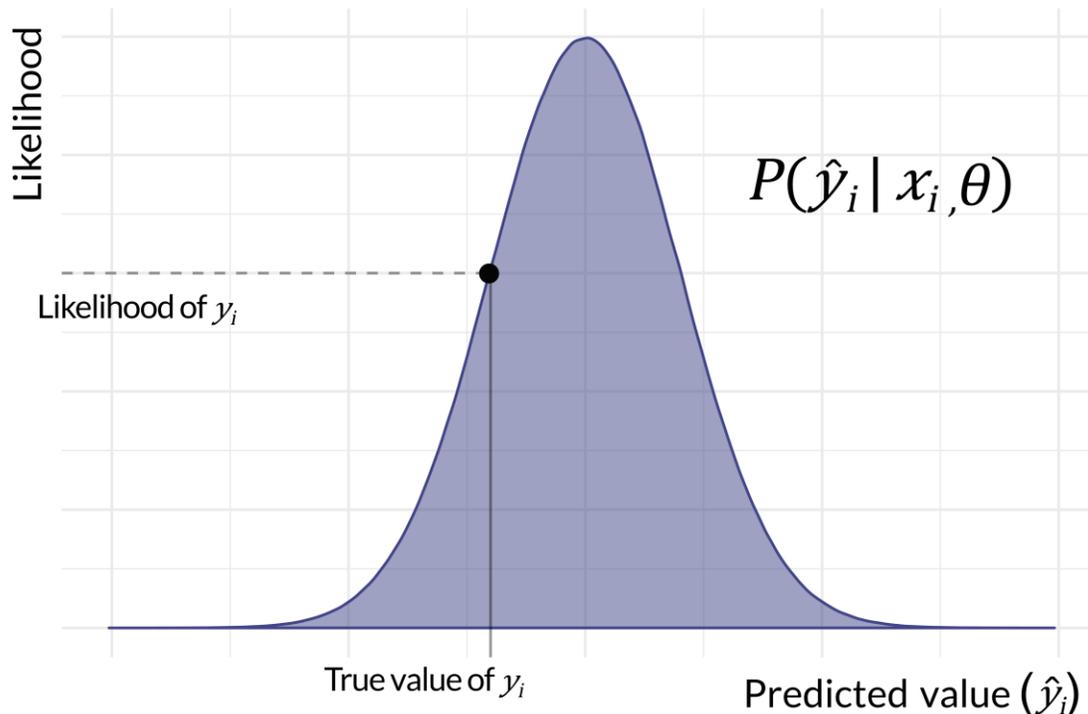


Figure 3.13: Posterior predictive distribution for a single observation x_i . The true value of y_i is plotted, and its likelihood ($p(y_i | \hat{y}_i)$) can be computed.

In summary:

- 1) A **model** is a mathematical representation of a real-world phenomenon. It describes the relationship between the characteristics of individuals that form a population.
- 2) Through the model, two entities interact: the **data** and the **parameters**. The data are the observed, measured quantities that we have collected. They are the **observations**, which are the individual of the population, described by a set of variables or **features**, which are the quantities measured. The parameters of a model are the unknown quantities that were not measured and need to be estimated.
- 3) Statistical **inference** refers to the process of estimating the real-world parameters from a sample of data. **Bayesian** inference aims to approximate the posterior distribution of the parameters, knowing the data.
- 4) **Point estimates** are single-value summary of a distribution. The mean, median and variance are all point estimates.

- 5) The **prior** is the initial probability of a parameter, before having collected the data. In applied Bayesian statistics, these can be based on initial beliefs of the statisticians. In science, this is rarely the case, and priors are used to constraint the values that a parameter can take (*e.g.* restrict a probability between 0 and 1) or regularize the model and reduce its variance.
- 6) The **likelihood** is the function that describes the data. It maps the observed data onto every possible values of parameters; it is the *plausibility* of the data for any combination of parameters, given the model.
- 7) The **posterior** is the conditional probability of a parameter given the data. This is the result of the Bayesian analysis.
- 8) **Posterior draws** or samples or **simulations** are the individual values that a parameter can have, numerically drawn from the estimated posterior distribution. In the text, the word “sample” is avoided because of the possible confusion with the brain tissue samples. Together, the posterior draws form the posterior distribution.

CHAPTER 4 ARTICLE 2: FEATURE ENGINEERING APPLIED TO INTRAOPERATIVE IN VIVO RAMAN SPECTROSCOPY SHEDS LIGHT ON MOLECULAR PROCESSES IN BRAIN CANCER: A RETROSPECTIVE STUDY OF 65 PATIENTS

The article presented in this chapter covers the core of the memoire. It describes the development of a new feature engineering method for Raman spectroscopy data that can make the spectra amenable to robust statistical analysis and unveil biologically meaningful patterns concealed in the Raman spectra.

This work was published in *Analyst* in October 2019 by Émile Lemoine with the help of postdoctoral student Frédérick Dallaire, R&D industrial engineers Rajeev Yadav and Rajeev Agarwal, pathologists Dominique Trudel and Marie-Christine Guiot, neurosurgeon Kevin Petrecca and engineers and professors Samuel Kadoury and Frédéric Leblond. The contribution of the first author includes the identification of the problematic, idea behind the proposed method, development and invention of the feature engineering method, data re-processing, computational experiments, interpretation and analysis of results, conception and creation of visualizations, and redaction of the manuscript.

This publication was reproduced with the permission of *Analyst* in the following pages. References are included at the end of the memoire.

4.1 Feature Engineering Applied to Intraoperative In Vivo Raman Spectroscopy Sheds Light on Molecular Processes in Brain Cancer: A Retrospective Study of 65 Patients

Journal: *Analyst*

Émile Lemoine^{1,2}, Frédérick Dallaire^{2,3}, Rajeev Yadav⁴, Rajeev Agarwal⁴, Samuel Kadoury³, Dominique Trudel², Marie-Christine Guiot⁵, Kevin Petrecca⁵, Frédéric Leblond^{1,2}

¹Department of Engineering Physics, Polytechnique Montreal, Montreal, Quebec, Canada

²Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montreal, Quebec, Canada

³Department of Computer and Software Engineering, Polytechnique Montreal, Montreal, Quebec, Canada

⁴ODS Medical Inc, Montreal, Quebec, Canada

⁵Brain Tumour Research Center, Montreal Neurological Institute and Hospital, Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada.

4.1.1 Abstract

Raman Spectroscopy is a promising tool for neurosurgical guidance and cancer research. Quantitative analysis of Raman signal from living tissue is, however, limited. Their molecular composition is convoluted and influenced by clinical factors, and access to data is limited. To ensure acceptance of this technology by clinicians and cancer scientists, we need to adapt the analytical methods to more closely model the Raman-generating process. Our objective is to use feature engineering to develop a new representation for spectral data specifically tailored for brain diagnosis that improves interpretability of the Raman signal while retaining enough information to accurately predict tissue content. The method consists in the band fitting of Raman bands which consistently appear in brain Raman literature, and the generation of new features representing the pairwise interaction between bands and the interaction between bands and patient age. Our technique was applied to a dataset of 547 *in situ* Raman spectra from 65 patients undergoing glioma resection. It showed superior predictive capacities to a Principal Component Analysis dimensionality reduction. After analysis through a Bayesian framework, we were able to identify the oncogenic processes that characterize glioma: increased nucleic acid content, overexpression of type IV collagen and shift in the primary metabolic engine. Our results demonstrate how this mathematical transformation of the Raman signal allows the first biological, statistically robust analysis of *in vivo* Raman spectra from brain tissue.

4.1.2 Introduction

Raman spectroscopy can recognize brain cancer cells *in vivo* during brain surgeries[44], [225]. Following recent integration in portable surgical tools[30], [42], [43], [45], [276]–[278], it holds significant potential to address a shortcoming in cancer research and diagnosis: the interrogation in real-time of human cancer cells *in situ*[8], [51].

With animals and *ex vivo* models, we have come a long way to identify molecular processes in brain cancer. In order to survive and proliferate, cancer cells adapt their behavior and environment in a highly complex but predictive way, making it possible to track down their presence using specific metabolic biomarkers. In glioma, the most prevalent form of brain cancer worldwide, these processes include, among others, neovascularization, hypoxic energy production, enhanced cell motility and replication, and genetic mutations[279], [280]. All of these perturb the tissue surrounding cancer cells and make glioma detection with Raman spectroscopy possible.

The Raman spectrum relays a signal pertaining to proteins, cell membranes, nucleic acids and lipids, which are straightforward to identify in a controlled setting where only few molecules interact, and where thousands of Raman acquisitions can be performed. In medical applications, however, both these conditions do not hold.

First, the Raman signal of living tissues is convoluted: every Raman band is not independent of the others[188]. One molecule can exhibit multiple peaks, and a single peak can belong to many molecules[159]. Brain composition is also strongly affected by individual patient characteristics. Among them, age is the most studied: older brains accumulate protein, lipid and carbohydrates deposits and undergo progressive genomic damages[281]. Second, published brain Raman studies include no more than 30 patients[225]. Access to human tissue is limited and time consuming and labeling of tissue by expert pathologists is costly. While Raman spectra can exhibit from 500 to 1,000 variables, the variables-to-observations ratio generally recommended to optimize the stability of statistical models ranges between 1:10 to 1:50[49], [282], [283].

Traditionally, analysis of biological Raman data abstracts away all this complexity by relying on machine learning models to mathematically learn the spectral patterns that map one spectrum to pre-defined labels (*e.g.* cancer tissue vs normal tissue)[10], [47], [160], [197], [203], [284]. Other methods have been proposed in an attempt to “unmix” spectral components from the brain Raman signal, most of which define the Raman spectrum as a linear combination of uncorrelated “endmembers” spectra[217]. While these algorithms may be convenient for specific applications (*e.g.* generating the three channels of a RGB image based on the Raman emission of a sample[115], [116], [176], [224]), they result in an oversimplification of the optical properties of the tissue by assuming that a small, finite number of components are present and that these components do not interact between each other. In addition, statistical robustness—susceptibility to small variations

in the data or model parameters—is rarely questioned, while uncertainties regarding tissue composition and signal composition is inherent to this technology[49], [164], [167]. Thus, although we know Raman can distinguish tissue phenotypes based on these approaches, there exists no method to robustly deconstruct and quantify the molecular composition of living tissues based on their Raman signal, limiting the power of Raman spectroscopy as a tool for *in situ* metabolic research.

A spectrum, however, is only one possible representation of the Raman signal. This “spectral” representation is convenient for visual assessment—but when automatizing the analysis of signals acquired in biomedical settings, its many pitfalls become evident: it contains a large amount of shot and biological noise, does not account for interactivity between bands, and cannot integrate clinical information about the tissue nor patient being tested[47]. Whereas most of the research on adapting Raman spectroscopy for *in vivo* diagnosis has focused on technical aspects of the systems, not much has been challenged in terms of analytical methods[22], [49], [188], and incorporating domain-knowledge into bio-spectroscopy analysis is poorly explored by fear of introducing subjective bias in a sterile statistical analysis[48]. These limitations hinder our ability to quantify molecular content of a sample from its Raman signature.

We present a new method to transform the Raman spectral information into a format that is interpretable and amenable to robust statistical analysis. We show how this method allows to account for the hierarchical and interactive structure of the Raman signal and control for clinical information by dramatically reducing the size of Raman data, all while preserving predictive power when compared to previous modeling strategies. We used this method to perform a retrospective, exploratory analysis of 547 human brain and glioma Raman acquisitions previously acquired *in situ* during 65 different neurosurgical resections and use a multivariable Bayesian statistical model to shed light on the cancer-specific mechanisms amenable to Raman systems in living brain cancer.

4.1.3 Methods

4.1.3.1 Spectroscopic data acquisition

Patients were selected by a single neurosurgeon (KP) during the years 2014 to 2018 at the neuro-oncology department of the Montreal Neurological Institute and Hospital (Quebec, Canada). In total, 135 brain cancer patients were operated while using a hand-held point-probe Raman

spectroscopy system, of which 65 suffered from glioma grade II to IV. The diagnosis of excluded patients included meningioma, lymphoma and metastases. Informed consent was obtained for all patients, and procedures were approved by the Montreal Neurological Institute and Hospital Ethics Review Board.

The Raman interrogations presented here occurred during the development of a hand-held Raman spectroscopy near-infrared point-probe designed to assist neurosurgeons in tumor resection procedures. During the 4 years of acquisitions, the systems including hardware and software evolved, the protocol slightly changed, and spectral data quality control became more rigorous. This time-dependent heterogeneity in the data is characteristic of any emerging technology and engenders considerable challenges when developing analytical solutions that depend on large datasets. These changes are presented in detail in **Supplementary Note 1**. Overall, they resulted in a higher proportion of high-quality spectra in the later years, with no significant changes in the Raman signal's general appearance throughout time (**Supplementary Figure 4.1**).

Data acquisition protocol and system specifications were similar in all patients and have been previously described[43], [44], [69]. The following terms are used to describe the experiment: a *Raman measurement* is a single capture of Raman scattered light, a *Raman acquisition* (or *Raman spectra*) is the ensemble of Raman measurements averaged for a single point, and a *brain sample* is the resected brain tissue corresponding to a single *Raman acquisition*. Briefly, the surgeon first planned the approach using pre-operative Magnetic Resonance Imaging (MRI), which he used to guide the surgery in combination with neuronavigation (Stealth Station; Medtronic). Raman acquisitions were performed in the trajectory of the surgery (an area that would be resected but appeared free of visible cancer to the operating surgeon), inside the tumor (as per the assessment of the surgeon informed by neuronavigation), or at the surgical margin after maximal resection. After every Raman acquisition, the interrogated area was sampled and fixed in paraffin for analysis by an expert neuropathologist. All Raman acquisitions consisted in excitation with a 785 nm light source (automated power control, 30-75 mW at the sample) followed by 3–10 50 ms Raman measurements (500 μm spot size). For all points, a 50 ms background measurement with no laser excitation was also performed. Multiple points could be acquired during a single surgery.

Many samples in the dataset were collected at the tumor margin, regions of high diagnostic uncertainty. While these samples can potentially lower statistical effects or machine learning

performances, they should not be discarded as they constitute the tissue most likely to be interrogated by the neurosurgeon during real-life procedures[49]. The molecular changes underlying glioma infiltration are still poorly understood and probably constitute a nonlinear continuum of processes that can hardly be divided under finite labels—some evidence suggesting that even nontumorous brain undergoes molecular changes in glioma patients[285]. However, pooling samples under distinct labels remains a simple way to lower the complexity of our statistical models and improve its stability. Moreover, less labels lead to more interpretable analysis. Therefore, we separated the samples under two labels defined prior to the statistical analysis: samples with no visible signs of glioma or with signs of infiltrative cancer that was non-diagnostic for glioma (*normal/low cancer density*), and samples with dense cancer (*high cancer density*). These labels were assigned based on H&E staining of the resected samples by an expert neuropathologist (MCG) blinded to the Raman analysis.

4.1.3.2 Signal processing and spectral quality

All raw signals, as detected by the CCD camera during acquisitions, were re-processed according to a state-of-the-art signal processing pipeline written in Python (version 3.6.5, 2018). First, the background signal (measurement acquired with laser turned off) was subtracted and the resulting spectrum was corrected for system response by normalizing to a measurement made on a Raman standard (SRM2241; NIST, Gaithersburg, Maryland)[286]. Afterwards, remaining low-frequency background (mainly tissue auto-fluorescence) was removed with the help of the Rolling Ball algorithm (ball width: 51 points)[287]. The spectrum was truncated to remove artifacts introduced by the Rolling Ball at both ends (150 points are removed on each side), for a final spectral range of 728 to 1730 cm^{-1} . A gaussian filter was applied. Finally, the spectrum was normalized with standard variate normalization (SNV) to have a mean of zero and standard deviation of one, and then the minimum was subtracted in order to keep all intensity values at or above zero. After signal processing, the full Raman spectrum contained 560 variables (**Supplementary Figure 4.2**).

Patient age was recorded for all acquisitions. Because of its bimodal distribution and to increase statistical power, the variable was dichotomized to “Young” and “Aged” at the median age of 52 years.

All spectra were also evaluated on their quality by three blinded, independent reviewers (EL, FD, GS) on the LabelBox platform (San Francisco, CA). Specific criteria were used such as: gross

signal-to-noise ratio, peak prominence and biological susceptibility of peak location (reference peaks were the phenylaniline peak at 1004 cm^{-1} , the nucleic acid peak at 1082 cm^{-1} , the amide III peak at 1300 cm^{-1} , the CH_2/CH_3 deformation peak at 1441 cm^{-1} and the amide I peak at 1659 cm^{-1}). Spectra were classified on a scale of 1–3 (higher values corresponding to higher quality), and the sum for all reviewers was kept as the final quality assessment. A score of 5 or superior was set as “higher-quality”. All individual spectra stratified by data quality score are presented in **Supplementary Figure 4.3**.

4.1.3.3 Raman feature engineering

The feature engineering method to change the representation of the Raman signal into one that is interpretable and amenable to statistical analysis involved two steps: 1) feature extraction with peak fitting and 2) mathematical modeling to create new features.

First, we extracted high-yield information from the processed Raman spectra to create a compressed representation of the signal that is easier to manipulate for the second step (**Figure 4.1A**). This is achieved through band fitting of pre-specified target bands. The target bands were selected a priori from the Raman literature, independently from our own data. This ensures no interference from our hardware, acquisition parameter or patient population, providing higher confidence of generalizability of the results outside our dataset. To select the target bands, we performed a systematic analysis of the Raman literature between 1999 and 2019, selecting all articles that presented original experiments on brain tissue[46], [49], [115]–[118], [169], [172]–[174], [176], [177], [179], [189], [190], [191, p. 1], [192], [200], [219], [223], [224], [288]. We excluded the articles in which the authors performed tissue or spectral manipulations that would result in significantly different spectra than our dataset. The exclusion criteria included paraffinization of the samples, different spectral range (*e.g.* high-wavenumber only), Coherent anti-Stokes Raman Spectroscopy, Surface-Enhanced Raman Spectroscopy, and articles from our own group. From the remaining 21 articles, we compiled every mention of Raman bands from the text or figures, and after accounting for band location variability ($\pm 5\text{ cm}^{-1}$), ranked the bands based on the frequency reporting. The 15 most frequently reported bands were selected as target peaks (as this was the median number of bands reported in the articles). For each of the 15 bands, in each spectrum, we identified the closest peak in a defined region around the band (5 wavenumbers on each side), found the first inflection points around the peak, and fitted a gaussian

density function to the spectra between those inflection points using Maximum Likelihood Estimation. Other functions can be used for band fitting (Lorentzian, Voigtian, Beta density function)[196]; the gaussian shape was selected as it resulted in a better fit for our spectra (**Supplementary Note 2**). Specifically, the gaussian function takes the form

$$\hat{y} = ae^{-\frac{(x-m)^2}{2s^2}}, \quad (4.1)$$

where \hat{y} is the estimated signal and x the x -axis values (in cm^{-1}), a is the maximal value or peak height, s is the standard deviation or the peak-width-at-half-maximum (hereinafter referred to as “peak width”), and m is the mean or band location (not used in further steps). If the estimation failed for one of the parameters, the values of all parameters were imputed: the maximum value of the spectra around the target band was used as height, and deviation of the intensities around the maximum as width (**Supplementary Note 2**). After this transformation, the data contained 30 variables: a peak height and peak width value for each of the 15 target bands.

Second, we created new features based on a defined mathematical model of the Raman signal in the tissue (**Figure 4.1B**). We used a binomial model to describe the probability that the interrogated sample contains highly dense cancer:

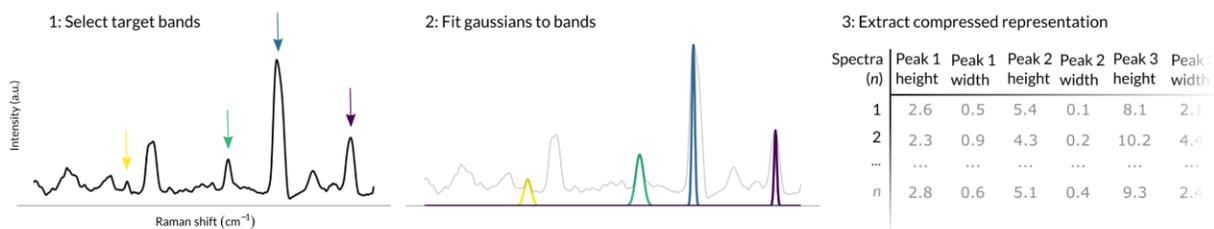
$$P(\text{High density cancer} \mid \text{Raman data}) \sim \text{binomial}(1, p \mid \text{Raman data}), \quad (4.2)$$

where p is the parameter of the binomial distribution $\binom{n}{k} p^k (1-p)^{n-k}$ for $n = 1$, where k is equal to 1 when the sample contains dense cancer and 0 otherwise. The value of p in eq. (4.2) can be expressed as a linear combination of variables. In our case, these variables included the peaks’ height and width and the patients’ age. The following model could be used:

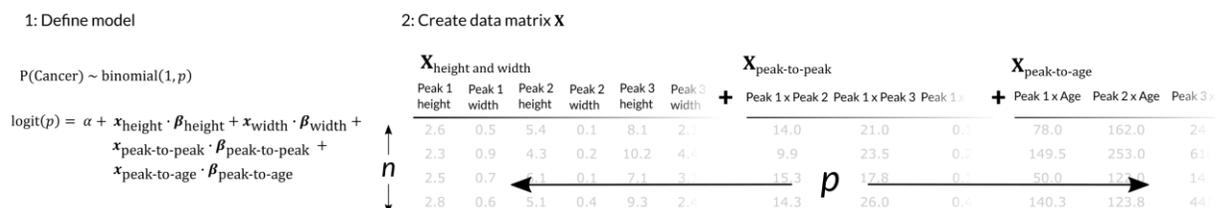
$$\text{logit}(p) = \alpha + x_{\text{height}} \cdot \beta_{\text{height}} + x_{\text{width}} \cdot \beta_{\text{width}} + x_{\text{Age}} \cdot \beta_{\text{Age}}, \quad (4.3)$$

where α is the intercept, β is a vector with the coefficients of the model we need to estimate and x contains the values extracted from the spectra or clinical record (the logit link restricts the output of the regression between 0–1). The model in eq. (4.3), however, treats every one of its parameters as an independent unit, whereas they should be considered as interdependent: the size of one peak affects how we should interpret the other peaks. We know normal brain and brain cancer composition varies with age. In the previous formula, age is considered an independent variable which directly influences the outcome (probability that the sample contains dense cancer), which

A - Feature extraction



B - Feature creation



C - Feature selection

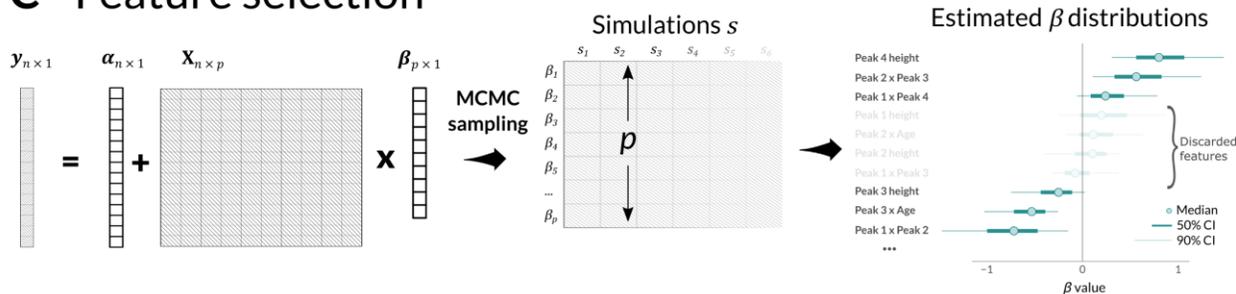


Figure 4.1 : Details of the feature engineering method. This is a conceptual figure; parameters are provided only to assist in visualization. A: Feature extraction steps. First, the target bands are selected (colored arrows). Second, these bands are fitted with a gaussian density function for each spectrum. Third, for each spectrum, the peak height and peak-width-at-half-maximum of the gaussian function is recorded. At this stage, the number of variables is equal to the number of peaks times two (height and width). B: Feature creation steps. First, a model (here, generalized linear model with logit link) is defined in order to represent the relationship between the Raman signal and the desired task (here, the detection of dense cancer in a brain tissue sample). Second, this model is applied to the representation issued from step A. At this point, the number of variables p has increased. C: Feature selection and inference. The data matrix X issued from step B and the labels vector y obtained from the neuropathological assessment of the samples are used to estimate the coefficient vector β in the model defined in step B. Each parameter has a corresponding coefficient β . The MCMC sampling finds S sets of solutions for each elements β_1, \dots, β_p of the vector β . The distributions for every coefficients β are then plotted, and each

coefficient is labeled with the feature it represents (rightmost panel). The distributions of coefficient for which the 50% highest density mass (thick line) encompasses zero are discarded (greyed-out).

is misleading; the patient age has no influence on the presence of cancer cells in the sample, but it does affect how we should interpret the Raman spectrum.

A Raman spectrum hides a hierarchy: bands cluster to form peaks, groups of peaks represent molecules and molecules interact to form structures or molecular pathways. This property of a variable to influence how another variable impacts the outcome is called an *interaction*, and every Raman band potentially has an interacting effect on the other Raman bands[188]. This means that the predictive effect of a Raman peak can only be considered in the context of other peaks—Raman bands are not independent. In the same way, we hypothesized that age has an interacting effect on the Raman signal. The generalized linear models typically used in multivariate analysis of spectroscopy data are agnostic to the spectral location of the variables; including interaction terms can overcome this limitation by forming a dependency network between the distinct spectral regions. Interaction effects can be modeled by multiplying the interacting variables together to create a new variable.

Adding interaction terms to a model improves its flexibility, but also increases exponentially the number of variables. With a limited number of observations, we needed to restrict the interaction terms to high-yield features that are most likely to explain the behavior of our system. In this case, we chose to put more confidence on two sets of interaction terms: *peak-to-peak* (pairwise interaction of a peak height value with all other peak height values) and *peak-to-age* (pairwise interaction between age and every peak height value). While it was retained as an independent predictor since we believed it could explain some of the variability in the height of the peaks, peak width was not used in interaction terms: we did not find strong evidence of its usefulness in biomedical Raman literature that could be worth the resulting increase in the number of variables. Therefore, we defined the following model instead of eq. (4.3):

$$\begin{aligned} \text{logit}(p) = & \alpha + \mathbf{x}_{\text{height}} \cdot \boldsymbol{\beta}_{\text{height}} + \mathbf{x}_{\text{width}} \cdot \boldsymbol{\beta}_{\text{width}} \\ & + \mathbf{x}_{\text{peak-to-peak}} \cdot \boldsymbol{\beta}_{\text{peak-to-peak}} \\ & + \mathbf{x}_{\text{peak-to-age}} \cdot \boldsymbol{\beta}_{\text{peak-to-age}} \end{aligned} \quad (4.4)$$

The terms on the first line in equation 4 are the variables for the peak height and width, and their coefficients describe the effect of these variables on the tissue content. The interaction terms between every pair of bands and between every band and the patient age are found on the second and third lines, respectively. Applying this model to our data leads to an expansion of the variables set from 30 variables to 150 variables (**Figure 4.1B**). This model constitutes a hypothetical mathematical representation of how the Raman signal explains tissue composition. This representation needs to be compared with other models used in the actual Raman literature in order to be validated. Then, we can use it to perform inference and answer more specific questions about how Raman signal is affected by tissue phenotype.

4.1.3.4 Statistical analysis

To estimate the coefficients of the model described by eq. (4.4) (and as such to estimate the quantified importance of every variable as predictor for dense glioma), we used a Bayesian framework (**Figure 4.1C**). Bayesian statistics rely on probability theory to estimate the likelihood of a state or event. *A priori* knowledge is updated by the data during the optimization process to recover the posterior probability distribution of a parameter. The *a priori* distribution can be used to restrict the complexity of a model, and acts as a feature selection filter. Here, we assumed that no extracted variable has a predictive value in identifying cancer in a sample by using an *a priori* distribution heavily centered at zero; this will result in a *sparse model*. In a sparse model, most parameters are equal to zero so that only truly important variables will impact the predictions. To draw the Bayesian optimization towards a sparse model, we used *hyper-lasso* priors on the β coefficients, as they were demonstrated superior to other regularizing priors in cases of high variables-to-observations ratio[264], [266], [289]. To recover the posterior distribution of each parameter, we used Markov Chain Monte Carlo (MCMC) sampling, more specifically a variant of the Hamiltonian Monte Carlo algorithm called “No-U-Turn-Sampling” (NUTS) as implemented in the Stan software[275]. The code for the statistical analysis was written in R and Stan.

Validation of the feature engineering pipeline described earlier consisted in comparison of its predictive performance with a Principal Component Analysis (PCA) dimensionality reduction method (as typically used in Raman studies[116], [173], [174], [290]). PCA projects the data into a set of linearly uncorrelated variables by computing the eigenvectors of the data’s covariance matrix, sorting them by the amount of explained variance (or eigenvalues). The centered data is

then projected onto these axes to obtain the Principal Components (PC)[291]. For this study, we reduced the data to the first 50 PC, a number that optimized the performances on our holdout set and expressed 85% of the variance in the data. The model comparison was performed by 10-fold cross-validation, where the dataset is split 10 times into a training set y_{train} and a holdout set y_{holdout} . Moreover, data from a single patient are kept together into either the training or holdout set to better replicate real-life predictions, where acquisitions from a patient will not be used to train the model that will perform prediction on that same patient. The parameters are estimated on the training set, yielding the “training predictions” $p(\beta_{\text{train}}|y_{\text{train}})$, and “testing predictions” are performed on the holdout set, resulting in a distribution of predictions $p(y_{\text{holdout}}|\beta_{\text{train}})$. Each prediction is done with a different set of coefficients β_{train}^s for each simulation s (or MCMC iteration s). To assess the predictive performance of the model on new data, we calculate its *log predictive density* (*lpd*):

$$lpd_{\text{holdout}} = \sum_{i=1}^n \log \left(\frac{1}{S} \sum_{s=1}^S p(y_i | \beta_{\text{train}}^s) \right), \quad (4.5)$$

where y_i is a single observation of the holdout set (of which every observation will be part of at some point during the cross-validation process) and S is the number of MCMC simulations (here: 8,000). The *lpd* reflects the accuracy of the predictions, but also the confidence in these predictions: a model that emits predictions highly concentrated around a certain value will be favored against a more conservative model for which predictions are more distributed on the entire prediction range. Three models were compared: 1) PCA with 50 PC followed by binomial model with normal (uninformative) priors, comparable to a PCA-Linear Discriminant Analysis (LDA) model, 2) PCA with 50 PC followed by a binomial model with hyper-lasso priors, and 3) the feature engineered Raman signal (band fitting followed by addition of peak-to-peak and peak-to-age interactions) followed by a binomial model with hyper-lasso priors. The objective of the model evaluation was to compare the different models—not to estimate the true performances of either one on unseen data.

Both PCA and gaussian fitting result in a compressed representation of the Raman spectra. After PCA, the original spectra can be reconstructed by multiplying the PC matrix (or scores) with the transposed matrix of the non-discarded eigenvectors (or weight associated with each feature). The amount of variance expressed in the discarded PC will affect the quality of this reconstruction. In

the same manner, the information from the gaussian fitting can be used to reconstruct the original spectra. To assess the information loss encountered during both feature reduction methods, we calculated the reconstruction error between the original (full) spectra, and the spectra reconstructed from both PCA and band fitting. The normalized root-mean-squared-error (nRMSE) was used as performance metric. We computed the nRMSE first on the full spectra and then on high-yield regions only (*i.e.* at the location of the target bands defined from the literature review).

To uncover the relative importance of each spectral band in discriminating normal/low-density and high-density glioma samples, we applied the Bayesian logistic regression with hyper-lasso priors to the entire transformed dataset (**Figure 4.1C**). A first run was performed with 4 chains and 4,000 iterations (1,000 burned-out) to diagnose convergence problems. Then, for inference, a single chain with 10,000 iterations (2,000 burned out) was used. To filter out uninformative variables, we selected those for which the estimated coefficient distribution was heavily centered at zero, *i.e.* having a large probability of being equal to zero. The threshold we set was that if zero was part of the 50% of values with highest probability, we considered the variable excluded from the model (**Figure 4.1C**, rightmost panel). In other words, only coefficients with a 75% probability mass above or below zero are considered non-zero.

4.1.4 Results

4.1.4.1 Description of patients and samples

Five hundred and forty-seven brain tumor samples from 65 distinct patients were interrogated by three different Raman probe systems during neurosurgical resection of glioma between 2014 and 2018. There was a median number of 6 points (and resected brain samples) per patient (range: 1—24). Of the 547 samples, 223 were selected as “higher-quality” by the evaluators (**Table 4.1**). Of these, 98 (44%) were labeled *high-density cancer*. The mean age of the patients was equivalent between both groups of tissue (50 years *vs.* 52 years). To our knowledge, this is the largest dataset of *in vivo*, neurosurgical human Raman spectroscopy (**Figure 4.2**).

4.1.4.2 Systematic literature analysis

Twenty-one articles were identified during the literature review, and all reported bands were recorded. In total, 61 bands coincided with the spectral range of our camera (median 14 bands per

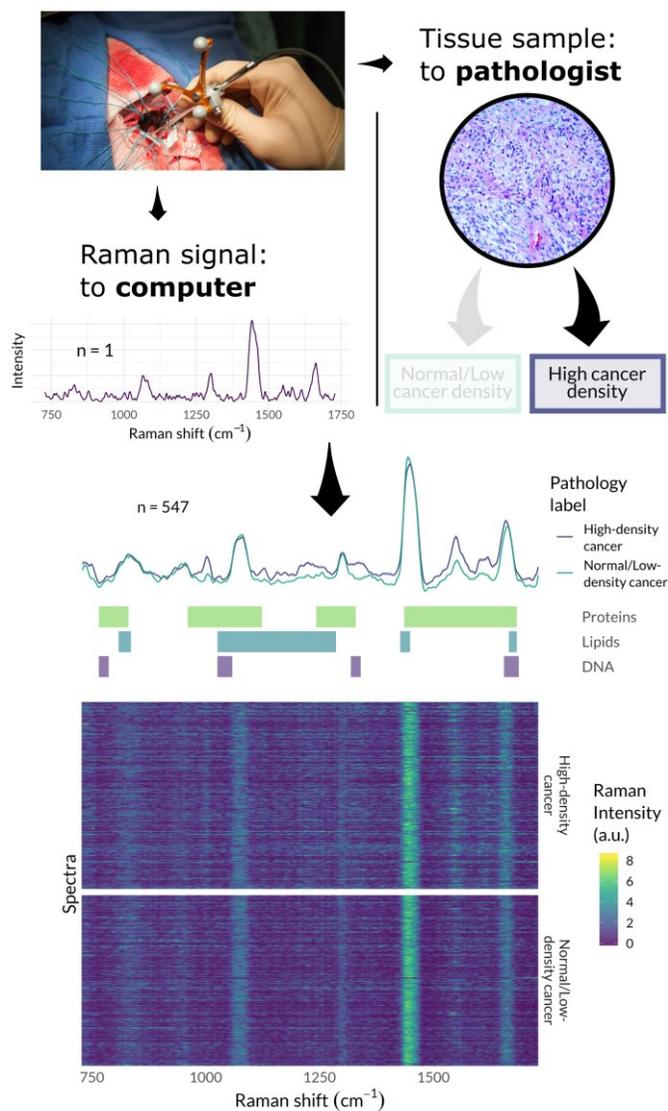


Figure 4.2 : Depiction of the workflow used for the acquisition of the Raman spectroscopy data.

The signal was acquired with a Raman handheld point-probe by a single neurosurgeon. After acquisition, the interrogated region was resected and processed for analysis by a neuropathologist. According to the degree of infiltration in the sample, the spectrum was labeled as ‘Normal/Low-density cancer’ or ‘High-density cancer’. The bottom heatmap represents all spectra included in the analysis (every line is a single Raman acquisition), color-coded for the Raman intensity. A.u.: Arbitrary units.

article, range 3 – 34) (**Supplementary Table 4.1 & Supplementary Figure 4.4**), and they all contained a Raman peak. The 15 most frequently reported bands constituted the final set of target bands for the Raman transformation routine (**Table 4.2**). The molecular assignment for each band is derived from previous Raman experiments on biological tissue, cells or fluids[159], [292]. Bands assigned to nucleic acid included 829, 1087, 1339 and 1659 cm^{-1} . Protein bands included amino acid markers such as 1004 and 1032 cm^{-1} (phenylalanine), 829, 852 and 877 cm^{-1} (hydroxyproline and tyrosine), and 1339 and 1553 cm^{-1} (tryptophan), amide bands such as 1659 cm^{-1} (amide I), and 1268 and 1300 (amide III), and a carotenoid band at 1158 cm^{-1} . Lipid bands were 877 (choline), 1087, 1268, 1128, 1064, 1300 (cholesterol), 1441 (fatty acids, cholesterol) and 1659 cm^{-1} (fatty acids).

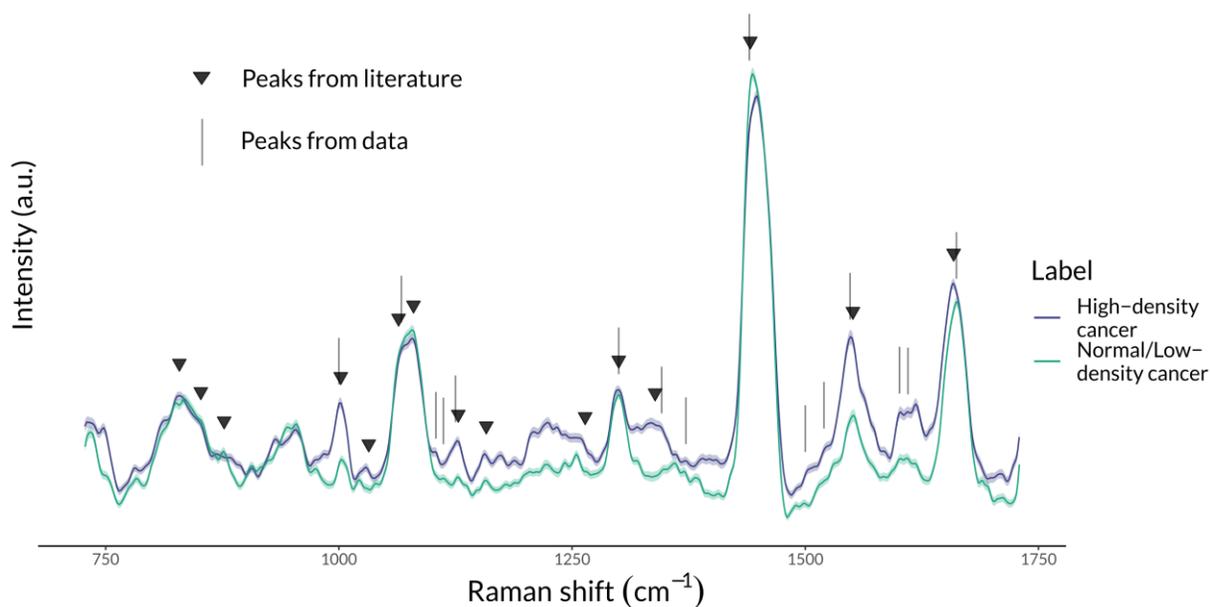


Figure 4.3 : Prevalent peaks in literature vs. prevalent peaks in the dataset. The average spectra for high-density and normal/low-density cancer are shown, with markers identifying the 15 target peaks selected after systematic review of the literature and the 15 most prevalent peaks found in the data with a peak finding algorithm.

A peak finding algorithm revealed the 15 most prevalent peaks present in our data. Of those, eight corresponded to the 15 literature bands (1004, 1064, 1128, 1300, 1339, 1441, 1551 and 1659 cm^{-1}). The target bands' location along with the location of the most prevalent peaks in our data are shown in **Figure 4.3**. All literature bands can be associated with a Raman peak in our data. The

seven other most prominent peaks in our data (1100, 1107, 1372, 1500, 1520, 1604 and 1610 cm^{-1}) were reported in respectively 5, 2, 3, 0, 6, 2 and 3 articles.

4.1.4.3 Extraction and creation of new Raman features

We applied the Raman feature engineering algorithm to the data. The spectra were first band-fitted using the 15 target bands recovered through the literature review (**Figure 4.4A**). The success rate of the gaussian-fitting was adequate (**Supplementary Note 1**).

We then generated the final feature set using the model defined in Methods, accounting for *peak-to-peak* and *peak-to-age* interaction effects. The final feature set contained 30 variables describing the peak width and height, 105 variables for peak-to-peak interactions, and 15 variables for peak-to-age interactions, for a total of 150 variables (**Figure 4.4B**). As a comparison, the same feature generation model applied to the whole spectrum (560 variables) would result in 157,640 variables, for a ratio of around 700 variables for every observation (*vs.* one variable for 10—50 observations as suggested by various authors[282], [283]). With a dataset of this size, considering interactions among peak and with clinical variables is only feasible with the highly condensed spectral representation achieved through the band fitting procedure.

4.1.4.4 Validation of the transformation method

Compressing the representation of the data undoubtedly results in information loss. To evaluate the impact of this loss on potential predictive performances, we compared the generated Raman features to a Principal Component Analysis (PCA) dimensionality reduction method, the standard in Raman spectroscopy for dimensionality reduction (**Figure 4.5A**)[48], [160], [188], [197]. Our objective is to demonstrate non-inferiority of the feature engineering pipeline compared to PCA for prediction on new data; as our approach to data representation optimizes biological interpretability of Raman data, equal predictive performances would still result in a superior approach over current practices.

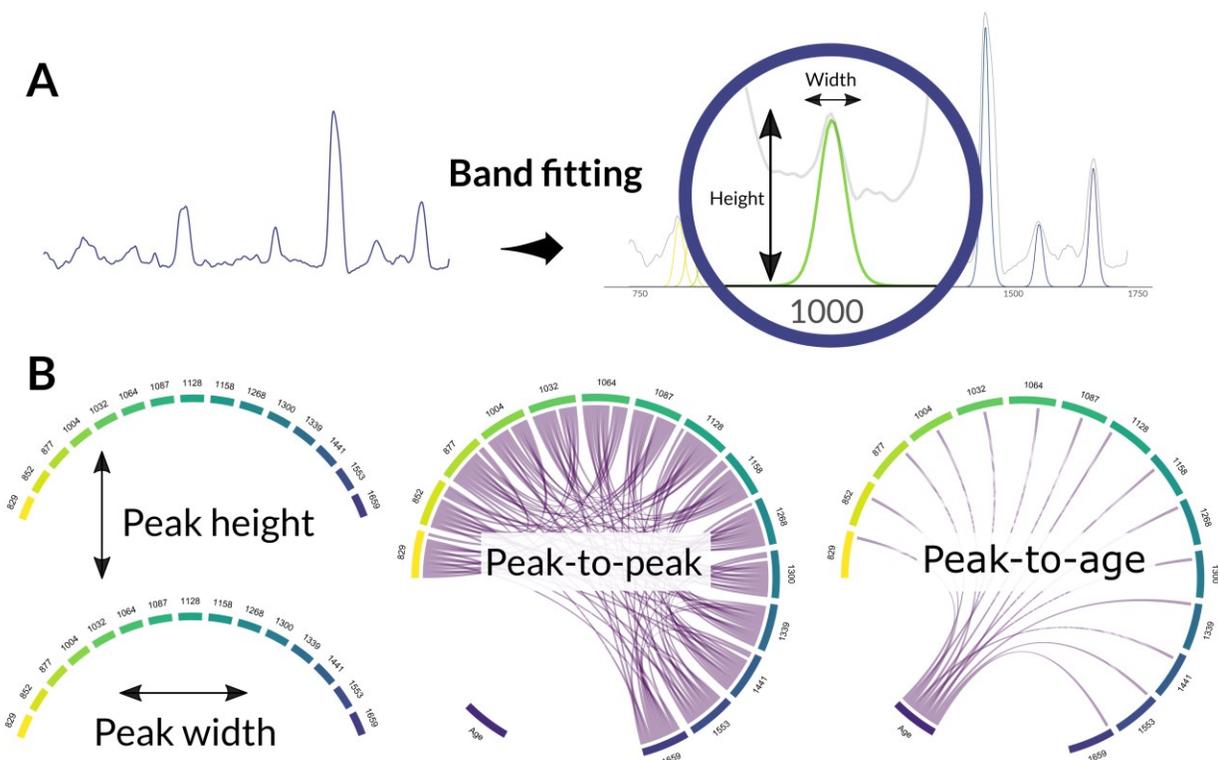


Figure 4.4 : Extraction of an interpretable representation for Raman data. **A:** The processed spectra is subjected to a band fitting algorithm. For each target band, a gaussian density function is fitted to the Raman spectrum, from which the height and width of the corresponding peak is estimated. **B:** The variables of interest that constitute the new representation. From left to right: peak height and width as estimated by the gaussian fitting, pairwise (peak-to-peak) interactions between every target bands, and interaction of age with every target bands (peak-to-age).

Under a Bayesian framework, we simulated real-life training and testing of a statistical model by means of cross-validation over 8,000 MCMC simulations (**Figure 4.5B** and **Figure 4.5C**). In 100.0% of the simulations, the band fitting feature reduction method was superior to PCA with normal priors (lpd with standard error : -160.1 ± 5.7 vs. -400.2 ± 218.3). In order to stabilize the predictive performances of the PCA method, we applied to it the same hyper-lasso priors than for the Raman transformation method. In that case, the band fitting was superior in 78.5% of the simulations (lpd with standard error : -160.1 ± 5.7 vs. -173.8 ± 17.2). Both results demonstrate the non-inferiority of the band fitting routine against PCA, and are suggestive of even higher predictive performances.

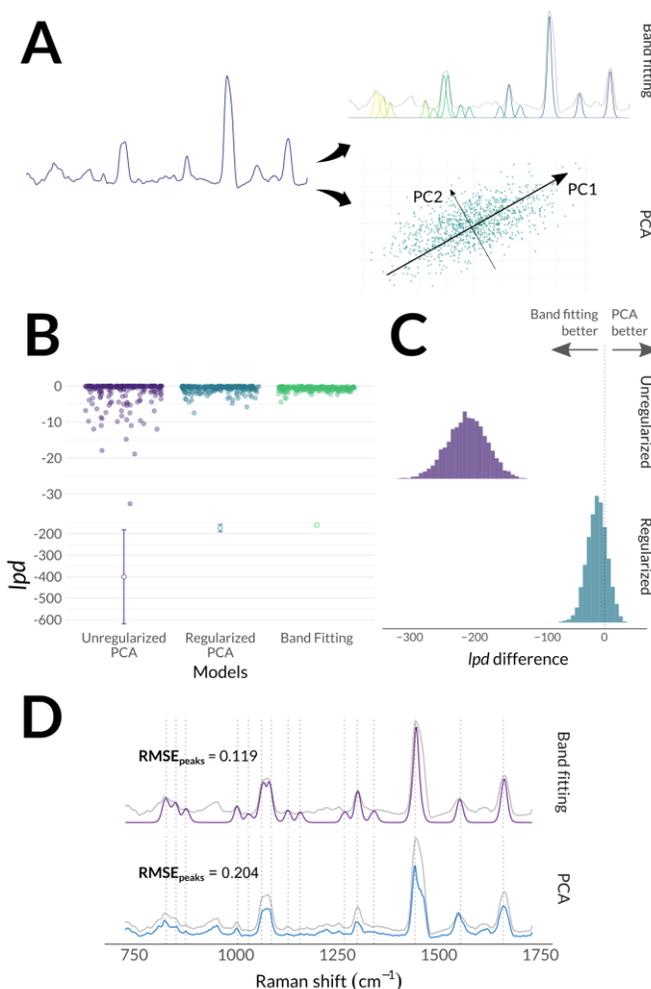


Figure 4.5 : Comparison of reconstruction error and cancer predictive performances between the Raman representation resulting from band fitting and dimensionality reduction following Principal Component Analysis (PCA). **A:** Processed Raman spectra were subjected to both band fitting (with subsequent addition of peak-to-peak and peak-to-age interactions) and PCA (50 principal components). **B:** lpd for every sample (upper part of figure) and cumulated over the full dataset (lower part) during 10-fold-cross-validation for the prediction of dense cancer in samples from neurosurgical tumor resection based on the Raman signature. A higher total lpd corresponds to higher predictive performances. **C:** Distribution of differences between the lpd of the band fitting representation and PCA representation, with and without regularization. Values under 0 represent a higher performance of the band fitting representation, and values over 0 a higher performance for PCA. For the unregularized PCA, 100% of simulations favored Band fitting, and for regularized PCA, 78.5% favored band fitting. **D:** Mean reconstructed spectra from both representations, with root-mean-squared error (RMSE) at target location.

We calculated the normalized RMSE (nRMSE) between the original data and the spectra reconstructed from both methods (PCA and gaussian-fitting)—a lowest nRMSE means a better fit. The nRMSE between the spectra reconstructed from the PCA and the original signal is 0.204, vs 0.522 for the band fitting. We are, however, more interested in how the reconstructions perform in brain-specific spectral regions. If we consider only the reconstruction error at the 15 target bands location, the nRMSE for PCA is 0.204, while it reduces to 0.119 for the band fitting (**Figure 4.5D**). While PCA performs better on the overall signal, the band fitting algorithm is more specific to regions that are highly likely to be of interest for our task (*i.e.* detecting cancer in a living brain).

4.1.4.5 Uncovering important features in Raman-based cancer detection

Next, we applied the Bayesian logistic regression with hyper-lasso priors to the entire transformed dataset. With the threshold defined in the Methods section, the model identified 30 variables as non-zero (**Figure 4.6A**). Important variables included peak heights, peak widths, peak-to-peak and peak-to-age interactions. To improve interpretability, variables were normalized according to their mean and standard deviation; when interpreting the coefficient of a single variable, all the other variables are set equal to zero (*i.e.* equal to the mean).

For seven bands, peak height had a high probability of association with the outcome (normal/low-density vs. high-density cancer) (**Figure 4.6B**). Five peaks were higher in dense glioma (1032 cm^{-1} [probability of this association: 78.4%], 1339 cm^{-1} [86.0%], 1004 cm^{-1} [89.0%], 1553 cm^{-1} [95.6%] and 1659 cm^{-1} [97.8%]), and two were higher in low cancer density samples (877 cm^{-1} [86.5%] and 852 cm^{-1} [99.1%]) (**Table 4.2**).

The 1004 and 1032 cm^{-1} bands are strongly associated with phenylalanine, and are frequently more intense in glioma samples [46], [49], [116]–[118], [169], [172]–[174], [179], [189]–[192], [200], [219], [224]. The 1553 cm^{-1} band is associated with the tryptophan amino acid and seen in glioma or necrotic tissue [116], [118], [169], [174], [179], [190], [191]. The 1339 and 1659 cm^{-1} bands are associated with multiple molecules. Strong 1339 cm^{-1} signal, a common marker of malignancy, is seen in proteins such as tryptophan, amide III and collagen but also is a prominent feature of nucleic acids [46], [115], [169], [174], [189], [190]. The 1659 cm^{-1} band is linked to amide I and proteins, lipids and nucleic acids, and has been reported both increased or decreased in glioma samples compared to normal tissue [116]–[118], [200], [219].

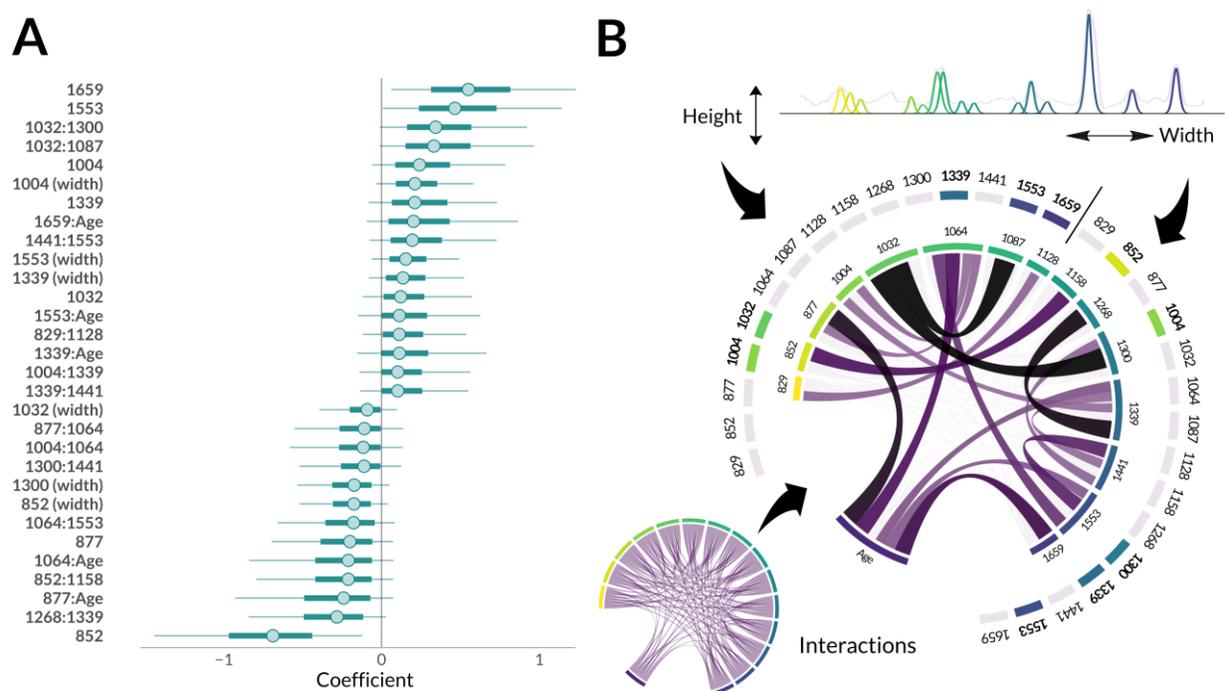


Figure 4.6 : Raman spectral features from the new representation selected as probably (>75%) informative of the presence of dense cancer in a sample by the multivariate analysis when controlling for peak width. **A**: Entire distributions of simulated coefficients for the features selected by the Bayesian model. Positive coefficients are associated with presence of dense cancer, while negative coefficients are associated with normal brain or low-density cancer. **B**: Graphical representation of the important features. Outer ring: colored bars represent peaks for which the height or width was selected as an important feature. Inner circle: dark links represent the interaction terms (peak-to-peak and peak-to-age) that were selected by the model. The width and darkness of the link correlates with the size of the probability (darker and larger links are more probable).

Normal or low-cancer density bands are related to tyrosine or proline compounds (877 and 852 cm^{-1}), and choline and cholesterol (877 cm^{-1}). These bands were previously reported as slightly higher in glioma samples or cells [172], [288].

4.1.4.6 Interactions and age

Of the 30 variables selected by the model, 11 were peak-to-peak interactions (**Figure 4.6B**). An interaction term represents a synergy between two bands: for example, these bands could be

expressed by the same molecule, or embody antagonistic mechanisms. Three bands appeared more frequently in peak-to-peak interaction terms: the 1064, 1339 and 1441 cm^{-1} bands. The 1064 cm^{-1} lipid band (phospholipids and cholesterol) had a negative effect on other bands (*i.e.* lowered their association with high-density glioma). It is usually reported as being less intense in glioma samples compared to normal brain [46], [49], [176], [179], [192], [219]. Both other bands were present in positive and negative interactions. They represent bands with strong overlap between molecular compounds (protein and nucleic acids for the 1339 cm^{-1} band, lipid and proteins for the 1441 cm^{-1} band), and both are usually reported as more intense in glioma samples [46], [49], [172], [189], [192], [200]. A few pairs of peaks denote an interaction between the concentrations of lipids and proteins (1004:1064 cm^{-1} , 1441:1553 cm^{-1} , 1339:1441 cm^{-1} , 1032:1087 cm^{-1})[159], [223], [293].

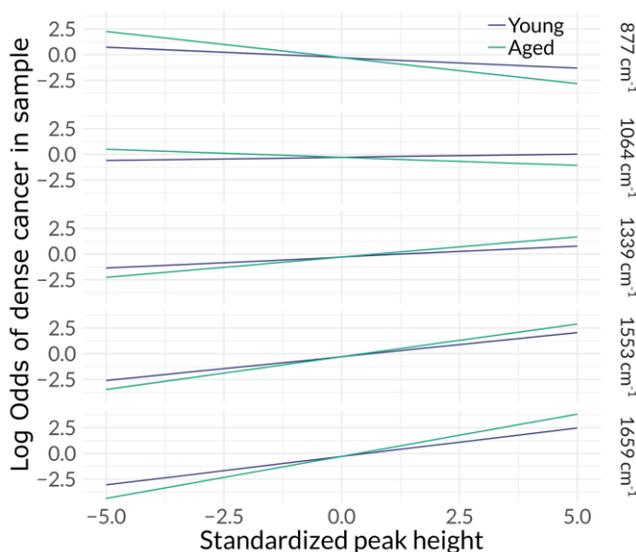


Figure 4.7 : Interactive effect of age on selected Raman bands. The slope of the association between the height of the peak and tissue type changes depending on the patient's age. If age was added as single variable, only the intercept would be affected.

The model selected 5 age-to-peak interactions from our data (**Figure 4.6B**). These peaks relate to the presence of nucleic acids (1339 [probability of association: 74.6%] and 1659 cm^{-1} [83.8%]), proteins (877 [87.0%], 1064 [86.1%] 1553 [75.3%] and 1659 cm^{-1}) and lipids (1659 cm^{-1})[159], [294]. **Figure 4.7** illustrates the effect of these interactions. In negative interactions (877 and 1064 cm^{-1}), and advanced age reduces the association between these peaks and the presence of high-density cancer. On the opposite, positive interactions (1339, 1553 and 1659 cm^{-1}) show how these spectral regions are more associated with malignancy for older patients than in younger patients.

4.1.5 Discussion

We proposed a new representation for spectroscopy data applied to Raman spectroscopy of *in situ* human glioma. This representation is specific to the Raman signal of the brain, accounts for the hierarchical structure in the data and allows the integration of clinical information such as patient age; we demonstrated how it is also more compact, amenable to statistical inference and minimizes information loss even with datasets of modest size as is typical in biomedical tissue experiments. This representation was applied to recover, in an exploratory and retrospective fashion, spectral features that appeared associated with the presence of dense glioma during neurosurgical tumor resection.

The tissue classification capabilities of Raman spectroscopy in neurosurgery has already been demonstrated[43], [44], [225], but their underlying molecular characteristics and mechanisms remain imprecise. The *in vivo* Raman literature fails to answer fundamental questions in this regard: which Raman bands are the most informative? Which Raman bands operate in unison? Does patient characteristics affect the Raman acquisitions? While this information is not necessary to predict tissue phenotypes in highly controlled experimental environments, it becomes essential when 1) expecting high predictive accuracy in real surgical situations, where acquisition time is limited and many confounding factors (ambient light, blood, probe manipulation, limited signal-to-noise ratio) can lower the confidence in the Raman data, and 2) performing chemometric analysis of *in situ*, infiltrative human cancer based on its Raman spectrum, which requires disentangling all biologically relevant information present in the data[22]. Attempts at band fitting and chemometric analysis of biological spectra have been avoided possibly due to the high heterogeneity of *in vivo* data and a fear of overly simplifying complex and convoluted molecular processes to the point of losing most of the information behind the molecular signature[48]. The spectral quality of our selected Raman spectra and the high reproducibility of important Raman bands between study groups made the reliance on band fitting possible, also ensuring Raman data was amenable to statistical analysis while preserving predictive power. Although there was some minor variability in experimental conditions between the acquisitions, the stability throughout time of the Raman signal in higher-quality spectra was essential to be able to pool a large quantity of data from an extended time period.

There is uncertainty when predicting the nature of a sample or estimating the effect of each variable; the 500 μm diameter interrogation integrates over a large variance in tissue content, the signal processing techniques involve direct manipulation and alteration of the signal, and acquisition parameter (state of tissue, lighting conditions, orientation of the probe) cannot be tightly controlled in a surgical context. Even sample labeling contains uncertainty, with samples harboring signs of cancer infiltration without being clearly diagnostic of glioma. The Bayesian framework used in this study allows to capture this uncertainty during both model evaluation and model interpretation, leading to a statistical inference that is more interpretable and nuanced. In the case of model comparison, not only can we identify the model with higher predictive performances, but also the likelihood of its predictions. When estimating the effect of each variable, we can assess these effects at different probability thresholds. Here, we filtered out variables with less than 75% probability of association with the outcome, a value that allowed a good balance between weeding out unhelpful variables while avoiding throwing away critical ones.

The model evaluation showed improved predictive performances for the feature engineering algorithm compared to PCA. It accurately retains Raman information from high-yield regions of the spectrum, namely those regions that are consistent across research groups, experimental settings and acquisition systems. Contrary to PCA, band fitting makes it possible to capture interactions between variable and integrate clinical information. Importantly, the band fitting procedure was not biased by the data at hand—all Raman spectra were processed the same, using *a priori* information available before the analysis. This information is derived from Raman expertise in brain tissue: using other Raman experiment as a source for our target bands stabilizes our result against variability inherent to our own data. Moreover, the fact that all the selected bands do show a prominent peak on our Raman spectra increases the confidence in the quality of the Raman signal in our dataset.

Targeting specific Raman regions allowed a simpler spectral representation, reducing the number of variables from 560 to 30. This step is mandatory to include interactions among peaks as well as interactions with clinical data, because of the quadratic increase in the number of variables; more specifically, for n peaks, peak-to-peak interactions yields $\frac{n(n-1)}{2}$ variables. Having variables more numerous than observations is commonly referred to as the *curse of dimensionality*[198]. In

Raman spectroscopy, this high-dimensionality results in many spectral bands being associated with the outcome only by chance, leading to unstable statistical models[47], [49], [255]. Regularization—purposefully limiting the flexibility of the model to avoid a close fit but improve generalizability to new data—allows to counter these problems[198]. Here, it involves two “filtering” steps. First, only reproducibly brain-specific bands, chosen without regards to the peaks visible in our data, were extracted during the band fitting, discarding probably confounding, artifactual or noisy information contained in other parts of the spectrum[295]. Second, hyper-lasso priors made most of the variables equal to zero at any point during the optimization of the Bayesian logistic regression. This allowed to reduce the variable count from 150 (a variables-to-observations ratio of 1:1.5) to 30 (ratio of 1:7.4), which is much closer to recommended ratios of 1:10–1:50 despite the relatively small size of the dataset. Regularization with Bayesian priors has previously been used in genetic studies—where variables-to-observation ratio can exceed 1,000:1—to improve accuracy and stability of analyses[296]–[298].

The major drawback of the hyper-lasso priors used here is the restriction in the initial number of variables—the higher this number, the lower the value of the final individual coefficients. While the model we defined in eq. (5.4) is plausible, its complexity is limited by the amount of data available. For instance, our results confirm the interest of peak-to-age and peak-to-peak interactions, but more complex interaction terms will need to be investigated with larger, independent datasets to ensure reproducibility of this model.

The model we proposed may be better than PCA at dimensionality reduction for Raman-based tissue prediction, but its true strength lies in its interpretability: the band fitting method allows to inquire the contribution of every variable to the model in a way that previous techniques could not. Although PCA supplies the relative amount of variance expressed by each variable, its use to identify important variables has severe limitations: when doing so, we assume that all variables are independent, and that variance is correlated with outcome[291]. However, most of the bands in Raman spectra contain high levels of shot noise (*i.e.* limited signal-to-noise ratio), or are consequences of phenomena not specific to the classification task at hand[160]. Also, the estimates provided by PCA (and other spectral “unmixing” methods) are point estimates—single values, without knowledge on the confidence of these estimates[116], [118], [169], [173], [174], [299]. On the other hand, as demonstrated, the data representation resulting from the feature engineering procedure targets task-specific spectral regions and recovers not only point estimates but a

distribution of estimates, making the data amenable to robust statistical inference. The recovered features could potentially be integrated into machine learning models to improve real-time cancer prediction to neurosurgeons during surgical resection; the performance of this method in terms of prediction accuracy would need to be estimated on a new, independent dataset.

The feature engineering method presented in this article allowed to recover spectral features that had a high probability of being associated with either normal/low-cancer density or high-cancer density during neurosurgical procedures. These engineered features have the advantage of carrying richer and more meaningful information than single bands, which rely heavily on the shape of the surrounding spectrum to be interpreted. There is growing concerns in the “-omics” fields to generate results and hypotheses that are not only robust but also functionally significant[20], especially since the application of this technology is targeted to non-Raman experts such as clinicians and cancer researchers.

Most of the cancer-associated features were markers of either nucleic acids or aromatic amino acids tryptophan and phenylalanine. Malignant tissue overexpresses nucleotides[279]. In glioma, this reflects cancer cell proliferation, reactive gliosis of surrounding tissue, or even the modification in the energetic engine of the tumor[279], [300]. Previous studies attempting to use Raman as a surrogate for DNA content in glioma have shown conflicting evidence, mostly since necrotic tissue has a low-DNA content but is prevalent in high-grade glioma samples [115], [116], [219]. Furthermore, the 1339 and 1659 cm^{-1} nucleic acid bands can also be assigned to other molecules (tryptophan, amide I and III, lipids). As such, while their significance in cancer detection is clear, their molecular interpretation remains uncertain.

Another possibility explaining the cancer-detecting capacity of the Raman probes rely on the upregulation of collagen by cancer cells, more specifically collagen type IV. In the brain, collagen type IV forms the extracellular matrices and lines the extraluminal vessel wall; both of which are upregulated in cancer[301], [302]. Indeed, malignant cells grow in a hypoxic environment and require an increased number of blood vessels to satisfy their increased metabolic needs. But this new vasculature is poorly efficient, and cells are forced out of the tumor via the extracellular matrices, which offer efficient migration tracts to a less toxic environment[303]. The cancer-associated bands located at 1004, 1032 and 1659 cm^{-1} have been shown as important markers for this compound[304], [305]. Moreover, many of the bands involved in interaction terms are also

related to collagen (829, 852, 877, 1004, 1032, 1064, 1158, 1268, 1339 and 1553 cm^{-1})[294], [304], [305]. In the past, strong focus has been put on the subtyping of lipids and lipid-associated molecules for characterizing malignancy in brain samples[215], [223]. Our results suggest that in the surgical setting, collagen may be a more prominent marker of malignancy for Raman-based systems.

Age showed interacting effect on bands associated with lipids, proteins and nucleic acids. The normal ageing brain accumulates protein, lipid and carbohydrates deposits[281], and undergoes significant genomic alterations that may be amenable to Raman spectroscopy such as methylation and DNA damages[306]. Furthermore, age can be a proxy for many conditions that affect brain constitution. IDH mutation, for example, is detectable by metabolic imaging and possibly by Raman spectroscopy[191], [307], and is highly correlated with age[232]. Our results do point towards the importance of including age as a controlling factor, but to validate and uncover the complexity of the underlying mechanisms would need to be addressed as a separate study.

DNA content, cell migration and adapted metabolic engine were the most prominent and probable manifestations of dense cancer in the Raman spectra, but other metabolic processes can also be derived from the cancer-associated bands. For example, we could attribute the importance of the 1659 and 1553 cm^{-1} bands to shifts in secondary structure of proteins, and 1659 and 1553 cm^{-1} to higher proportion of tryptophan[159], [294], [308]. These proposed mechanisms appear inconsistently in Raman literature. For secondary structure, some experiments demonstrate a shift from α -helix to β -sheets[118], [309], while others show the opposite[310], [311]. In the case of tryptophan, Raman studies (including ours) often indicate an increase in its relative concentration[310], while recent oncology research agrees that its depletion promotes immunosuppression by cancer tissue[312]. Also, our results demonstrate a strong association between the 852 and 877 cm^{-1} peaks and normal or low-density cancer, while these bands were noted significant in only two previous studies, where they were associated with malignant cells[172], [288]. These discrepancies might first stem from the fact that other studies used univariate statistical tools (or visual assessment only) to compare spectra, which can show erroneous correlations that would disappear after statistically correcting for the shape at other regions of the Raman spectrum[242]. Effect of small patient population could also be in cause but is difficult to estimate because previous works do not provide confidence estimation. In our analysis, limitations include inclusion of infiltrative cancer samples, spatial resolution of the

Raman systems and heterogeneity between glioma subtypes. Providing more data and adequate statistical power, a more complex model could be designed to stratify for glioma stage.

As hypothesized, peak-to-peak and peak-to-age interactions contributed heavily to the model. Caution must be taken when interpreting the results of a statistical analysis with high variables-to-observations ratio; still, after regularization, some of these features continued to exhibit a high probability of being important. The interdependencies of spectral features is inherent to the measurement we acquire, and while this study shines a light on their importance, the study of peak interactions should be addressed on its own to adequately reveal higher-order interactions (*e.g.* clusters of more than two peaks that vary together), and move from a one-dimensional representation of a spectrum to a hierarchical representation such as the one presented here. A similar shift in paradigm was observed in image analysis, with the introduction of deep, hierarchical architectures such as convolutional neural networks[313], and graph-network representations are emerging as a powerful modeling strategy for Big Data problems in biology[314]. Also, the effect of age on Raman spectra, as well as the impact of other clinical factors such as gender and ethnicity should be studied prospectively. Understanding the effect of those important confounders on the Raman data (and all spectroscopy data for that matter) will be an important breakthrough for the clinical translation of predictive models into medical practice.

4.1.6 Conclusion

This paper presents a technique to improve the representation of Raman spectroscopy data for the *in vivo*, non-destructive molecular profiling of human glioma. The proposed representation allows to better model the convoluted Raman-generating process; it focuses on high-yield regions of the spectrum, controls for the complex interactions between spectral features and allows to incorporate clinical data into the analysis. Our results demonstrated how relevant these factors are to predict the presence of cancer in target tissue, and that their implication should not be overlooked in the implementation of Raman-based surgical devices. Optical technology offers a unique opportunity to interrogate diseases in their own environment—their clinical translation will require data analysis methods adapted to the complexity of the world they try to model.

4.1.7 Conflicts of Interest

K. P. and F. L., are co-founders of ODS Medical Inc, a medical device company that seeks to commercialize the Raman spectroscopy system for real time detection of tissue abnormalities.

4.1.8 Acknowledgements

This work is supported by the Discovery Grant program from Natural Sciences and Engineering Research Council of Canada (NSERC), the Collaborative Health Research Program (CIHR and NSERC), the Canada First Research Excellence Fund through the TransMedTech Institute and MITACS. D.T. receives salary supports from the Fonds de Recherche du Québec–Santé.

4.1.9 Tables

Table 4.1 : Description of selected, high-quality tissue samples used for the Raman acquisitions

	High cancer density	Normal/Low cancer density	P-value
n	98	125	-
Age (mean (sd))	50 (16)	52 (15)	0.508
Gender = M (%)	46 (47)	69 (55)	0.276
IDH = Wildtype (%)	52 (53)	68 (54)	0.949
WHO Grade (%)			0.091
2	14 (14)	25 (20)	-
3	24 (24)	17 (14)	-
4	60 (61)	83 (66)	-

Table 4.2 : Raman bands selected as target for the band fitting procedure, with molecular assignment based on literature. The association with tissue phenotypes is based on findings from the current study.

Raman band (cm ⁻¹)	Associated bond	Assigned molecule	Molecular family	Association with tissue phenotype
829	O-P-O stretch PO2- stretch	Nucleic acids Tyrosine Proline	DNA Proteins	None
852	Ring breathing	Tyrosine Proline	Proteins	Normal/Low cancer density
877	C-C-N symmetric stretching C-O-C ring	Choline Tyrosine Hydroxyproline Sphingomyelin	Lipids Proteins	Normal/Low cancer density
1004	Symmetric ring breathing vC-C	Phenylalanine Collagen Heme Carotenoid	Proteins	High cancer density
1032	C-H CH2CH3 bending	Phenylalanine Collagen	Proteins	High cancer density
1064	C-O stretch C-O-C symmetric stretch C-C stretch	Proline Phospholipids side chains Cholesterol	Proteins Lipids	None
1087	C-C stretch PO2- symmetric stretch C=O vibration	Phospholipids Nucleic acids	Lipids DNA	None
1128	C-C stretch C-N stretch CH2 vibration	Phospholipids Cholesterol Glucose	Lipids Glucose	None
1158	C-C C-N	Collagen Carotenoid	Proteins	None
1268	Delta=CH Amide III	Unsaturated fatty acids Collagen	Lipids Proteins	None
1300	CH2 twist and wag Amide III	Phospholipids Palmitic acid Cholesterol Collagen	Lipids Proteins	None
1339	CH2 wag Ring breathing mode (DNA)	Nucleic acids Aliphatic amino acids Tryptophan	DNA Proteins Glycogen	High cancer density
1441	CH2/CH3 deformation	Lipids side chains Amino acids Cholesterol Cholesterol ester Collagen	Lipids Proteins	None
1553	vC=C	Porphyrin Tryptophan	Proteins	High cancer density
1659	Amide I vC=C	Nucleic acids Collagen Unsaturated fatty acids	DNA Proteins Lipids	High cancer density

4.1.10 Supplementary information

4.1.10.1 Supplementary Note 1

Acquisition parameters and experimental conditions per time period:

2016: A fiber-optic probe produced by Emvision, LLC that allowed intrinsic fluorescent spectroscopy (IFS), diffuse reflectance spectroscopy (DRS) and Raman spectroscopy (RS) was used. The probe is connected to an illumination and detection assembly using 3 m long fiber-optic cable with 2.1 mm diameter. The cable and probe consist in nine low hydroxyl content 300 μm core silica fibers (of which seven are used to collect Raman scattered photon) and notch filter for Raman collection, and a central 272 μm core silica fiber with an inline short-pass filter for Raman excitation. A two-component lens at the tip of the probe ensures spatial correlation between RS and other modalities. The seven RS collection fibers combine into one optical connector and are connected to a spectrograph composed of a diffraction grating and high-resolution charge-coupled device (CCD) camera (ANDOR Technology). Excitation is provided by a 785 nm spectrum stabilized laser (Innovative Photonic Solutions)[44].

During this period, sequential acquisition of DRS, IFS and RS was performed on all samples. Patient population included 7 patients with stage 2–4 glioma, for a total of 70 brain samples. The acquisition protocol is described in the main text. For RS, one 50 ms background measurement (laser turned off) and three 50 ms RS measurements were performed on each sample. The output power at the end of the probe varied from 27 to 75 mW to maximize gain.

2017 (SR): A fiber-optic probe produced by Emvision LLC for RS was used. The probe consists in seven low-hydroxyl, 300 μm core silica fibers and a notch filter, surrounding a central 272 μm laser-delivery fiber with a small band-pass filter. The front lens consists in a Plano convex 2 mm diameter curvature sapphire black portion a flat front and a of 1mm thick Plano magnesium fluoride to limit interference from the sapphire while allowing overlap of the focus from the illuminated and collected lights. The probe's collection system and its connection are identical as the one described above. The exception is the use of a[69]

During this period, the probe was used solely for RS. The configuration allowed to interrogate Raman signal for both the fingerprint and the high wavenumber spectral range. The population includes 4 patients undergoing neurosurgery for tumor resection of a glioma grade 2–4, for a total

of 41 brain samples. Again, the acquisition protocol and parameter were identical as the previous study period.

2018 (ODS): The same probe and collection systems were used.

The patient population included 22 patients undergoing neurosurgical resection of a glioma grade 2–4, for a total of 276 brain samples. The acquisition protocol and parameter were identical as the previous study period, with the exception that every Raman measurement consisted in one 50 ms background measurement (laser turned off) and ten 50 ms RS measurements

2019 (TRIAL): The same probe and collection systems were used.

The patient population included 32 patients undergoing neurosurgical resection of a glioma grade 2–4, for a total of 160 brain samples. The acquisition protocol and parameter were identical as the previous study period, with the exception that every Raman measurement consisted in one 50 ms background measurement (laser turned off) and five 50 ms RS measurements.

The resulting spectra for each period are illustrated in **Supplementary Figure 4.1**

4.1.10.2 Supplementary Note 2

The fitting of the Raman peaks estimates three parameters: the peak location, width (more precisely, full-width-at-half-maximum), and height. The parameter estimation can fail at three points: 1) peak maximum search, 2) optimization of the fitting procedure and 3) estimation of the parameters for the gaussian distribution. First, if no local maximum is found in a range of ± 5 cm^{-1} , the band with the maximum intensity in this range is used as the peak location and the gaussian function is fitted accordingly. Second, the gaussian fitting optimization could failed if it did not find an optimal solution after 1,000 iterations. Third, if the fitting was successful, the estimated parameters were compared to a range of “plausible” values according to the location of the target band and height and standard deviation of the signal at that location. If the optimization failed or any of the estimated parameters was flagged as erroneous, the values of all parameters were imputed: the maximum value of the spectra around the band was used as height, and deviation of the intensities around the maximum as width. Gaussian and Lorentzian densities were tested: fitting with Lorentzian function resulted in a normalized Root-mean-squared-error (nRMSE) at target band location of 0.128, while the nRMSE reduced to 0.119 with gaussian shapes. Gaussian shapes were used throughout the study.

Over all observations ($n = 547$), the variance across all spectra for each parameter was highly correlated with the quality assessment of the spectra ($F_{4, 24610} = 47.19$, $p < 0.0001$). Post-hoc analyses revealed that the estimated height and location of the peaks showed higher variance in low quality samples. This motivated the removal of poor-quality observations from the dataset; only higher-quality observations are considered ($n = 223$).

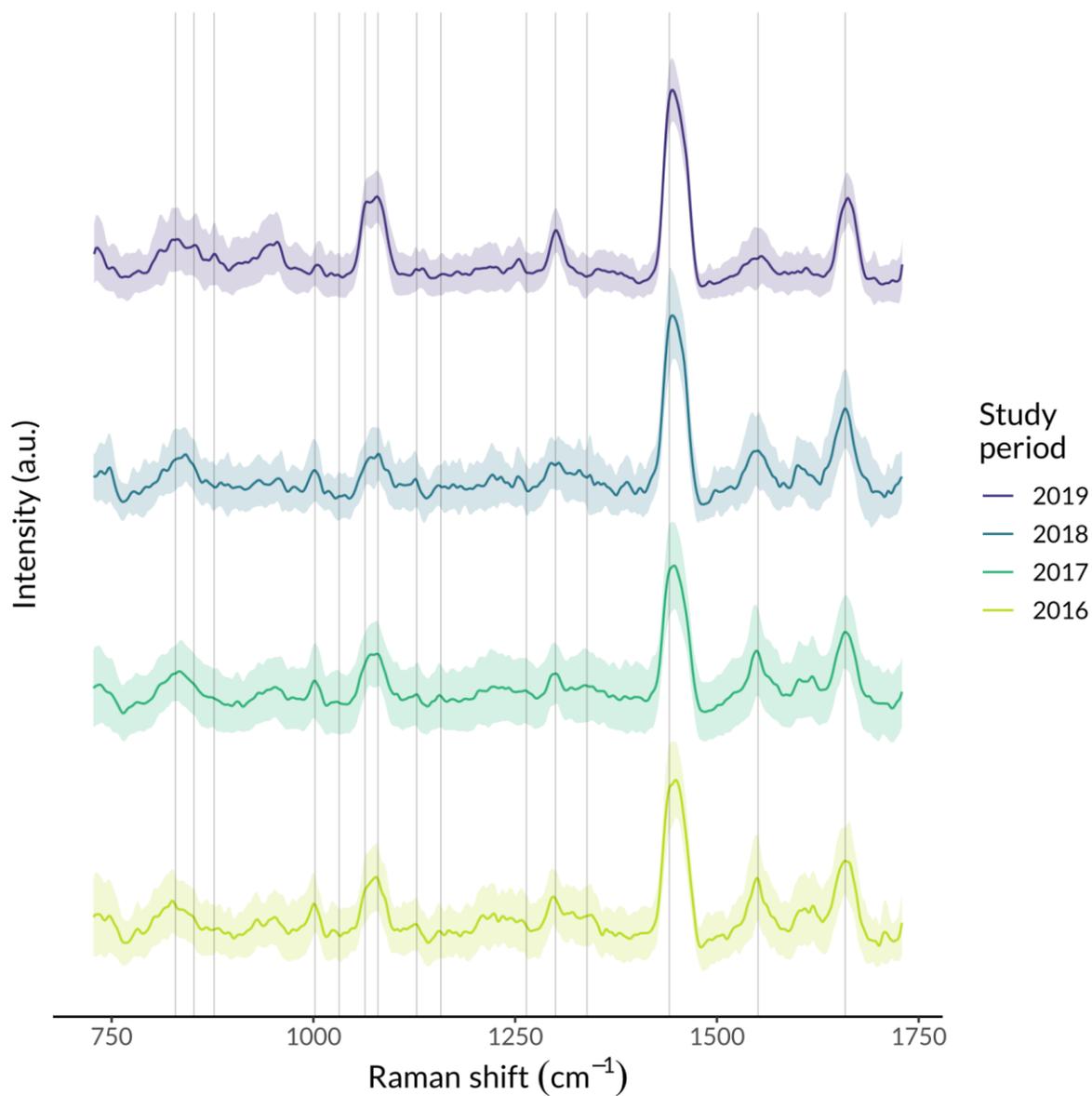
To assess the quality of the band fitting, we calculated the proportions of successful parameter estimation in the high-quality observations (15 target peaks over 223 spectra, for a total of 3,345 band fitting procedures). For 81.2% of the 3,345 peaks, all three parameters were successfully estimated. In all, 91.3% of total parameter estimations were successful. There is a slight variation of the fit quality between peaks: the bands at 1339, 1268 and 1032 cm^{-1} had the lowest rate of correct parameter estimation (81.6%, 84.3% and 85.7% respectively).

4.1.10.3 Supplementary Tables

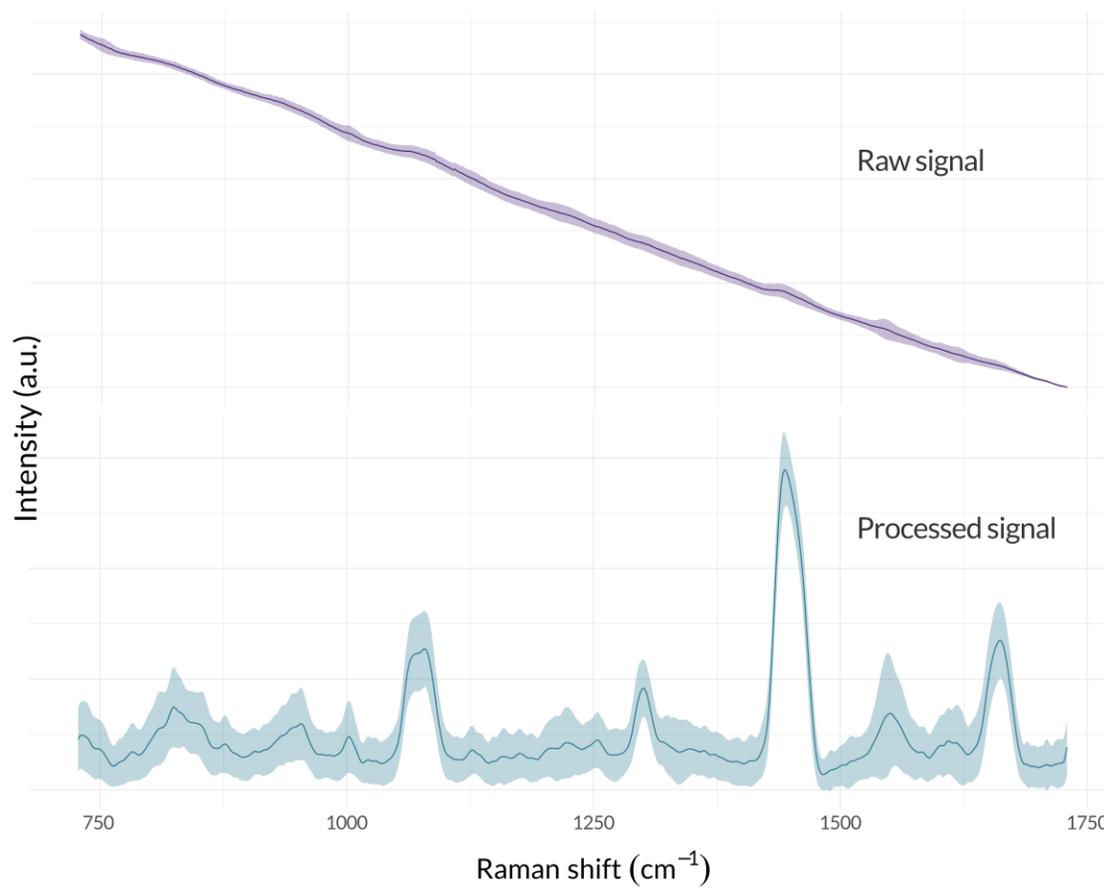
Supplementary Table 4.1 : Raman peaks reported in literature on brain tissue from 1999 to 2019.

Peak	N reported	Percentage reported	Raman shift range (cm ⁻¹)
1659	20	90.9	[1657 - 1661]
1441	19	86.4	[1436 - 1447]
1004	18	81.8	[1002 - 1008]
1300	17	77.3	[1296 - 1305]
1064	15	68.2	[1061 - 1068]
1339	14	63.6	[1333 - 1345]
1128	13	59.1	[1122 - 1132]
1268	13	59.1	[1262 - 1270]
829	9	40.9	[826 - 835]
852	9	40.9	[850 - 854]
877	9	40.9	[873 - 881]
1032	9	40.9	[1030 - 1035]
1087	9	40.9	[1081 - 1090]
1158	9	40.9	[1154 - 1161]
1553	9	40.9	[1550 - 1559]
785	8	36.4	[780 - 788]
936	8	36.4	[934 - 940]
1251	8	36.4	[1247 - 1255]
1175	7	31.8	[1172 - 1179]
1208	7	31.8	[1206 - 1212]
1671	7	31.8	[1669 - 1673]
757	6	27.3	[754 - 760]
1449	6	27.3	[1446 - 1453]
1523	6	27.3	[1518 - 1527]
925	5	22.7	[921 - 927]
1096	5	22.7	[1093 - 1100]
1583	5	22.7	[1581 - 1585]
1731	5	22.7	[1725 - 1735]
960	4	18.2	[958 - 962]
1315	4	18.2	[1313 - 1318]
1633	4	18.2	[1629 - 1635]
1742	4	18.2	[1739 - 1746]
1240	3	13.6	[1239 - 1240]
1375	3	13.6	[1374 - 1376]
1420	3	13.6	[1419 - 1422]
1563	3	13.6	[1560 - 1566]
1613	3	13.6	[1610 - 1614]
1656	3	13.6	[1655 - 1657]
748	2	9.1	[746 - 751]
898	2	9.1	[896 - 900]
1106	2	9.1	[1104 - 1107]
1226	2	9.1	[1225 - 1226]
1260	2	9.1	[1259 - 1260]
1275	2	9.1	[1272 - 1278]
1487	2	9.1	[1486 - 1488]
1577	2	9.1	[1576 - 1578]
1604	2	9.1	[1604 - 1604]
1668	2	9.1	[1668 - 1668]
739	1	4.5	[739 - 739]
845	1	4.5	[845 - 845]
865	1	4.5	[865 - 865]
954	1	4.5	[954 - 954]
977	1	4.5	[977 - 977]
1133	1	4.5	[1133 - 1133]
1199	1	4.5	[1199 - 1199]
1255	1	4.5	[1255 - 1255]
1295	1	4.5	[1295 - 1295]
1397	1	4.5	[1397 - 1397]
1461	1	4.5	[1461 - 1461]
1616	1	4.5	[1616 - 1616]
1623	1	4.5	[1623 - 1623]

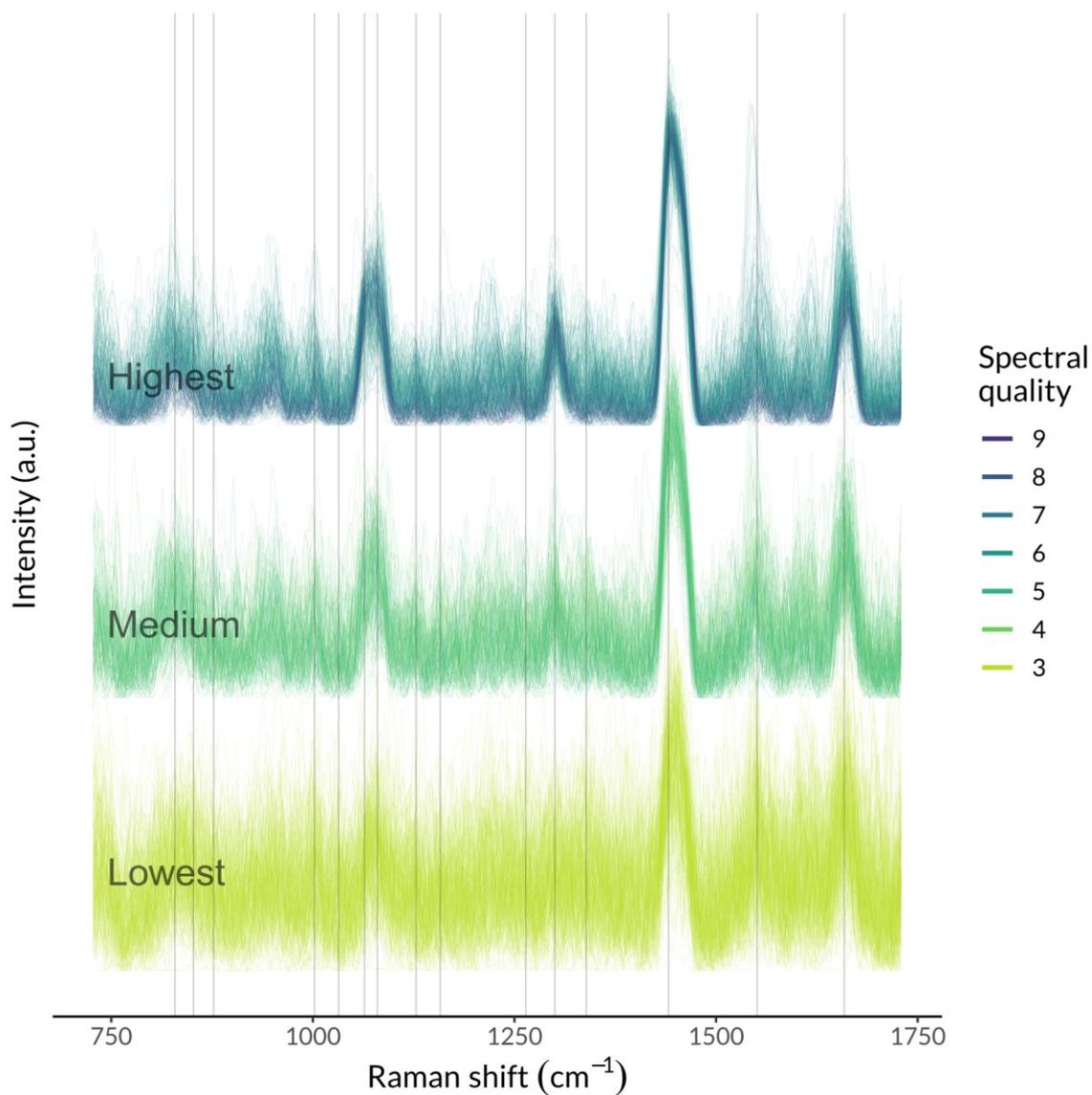
4.1.10.4 Supplementary Figures



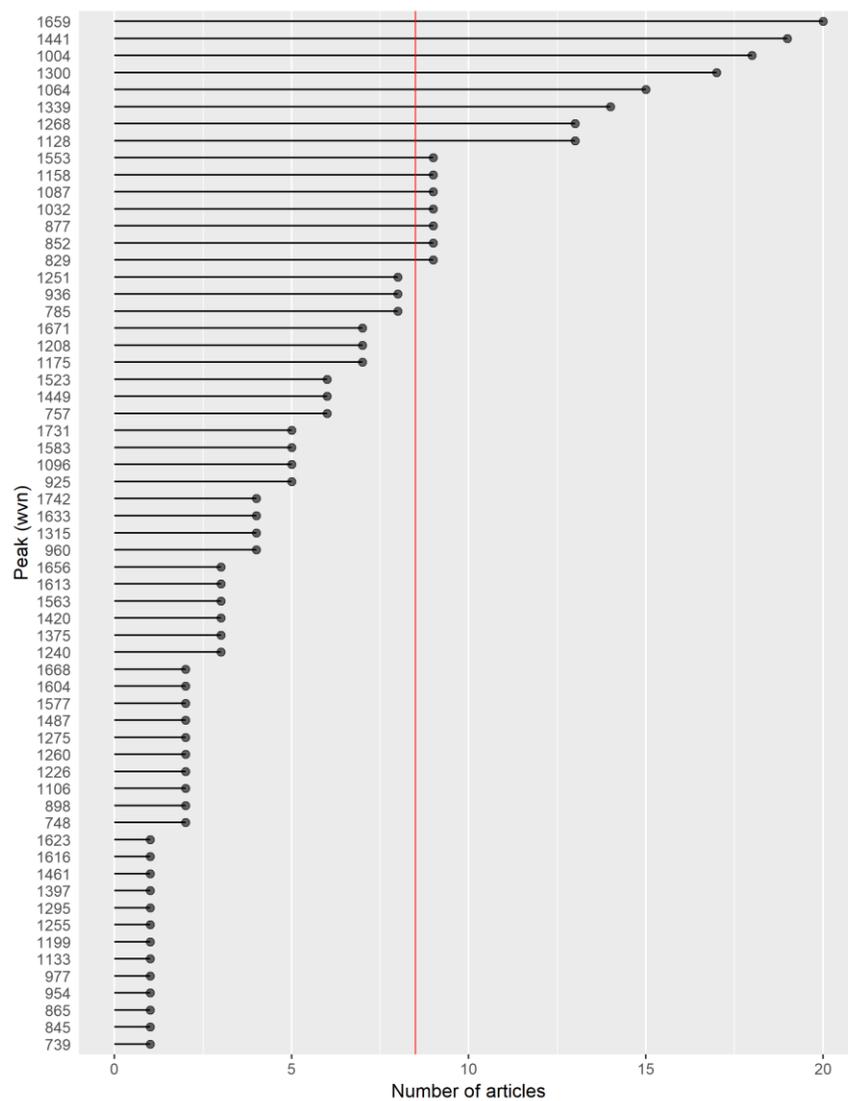
Supplementary Figure 4.1 : Averaged Raman spectra with standard deviation acquired in human brain samples during neurosurgical glioma resection, stratified by acquisition time periods. The vertical lines represent the 15 target bands used in the statistical analysis.



Supplementary Figure 4.2 : Averaged Raman acquisitions before (“Raw spectra”) and after data processing.



Supplementary Figure 4.3 : All individual Raman acquisitions ($n = 547$) stratified by data quality. Data quality was assessed by three independent reviewers with a rating of 1 (lower quality) to 3 (higher quality). The final score is the sum of all three ratings. A score of 5 or higher (“Highest” in the figure) was set as cut-off before statistical analysis. The vertical lines represent the 15 target bands used in the statistical analysis.



Supplementary Figure 4.4 : All Raman bands mentioned in the literature on Raman spectroscopy of brain tissue from 1999–2019. The red line indicates the cut-off for bands that were selected as target bands in the study.

CHAPTER 5 GENERAL DISCUSSION

This memoire describes the development of a feature engineering pipeline to extract molecular significance from intraoperative Raman spectroscopy. The pipeline comprises of a series of steps that improve the representation of human Raman spectra for statistical learning. Spectral features are first extracted from the Raman data and combined in a clinically relevant model, allowing integration of prior domain-specific expertise. Then, the most important features are selected via a Bayesian regularized logistic regression with multilevel priors, a state-of-the-art, probabilistic approach to separate informative features from the noisy variables. To test its validity, the pipeline is compared to Principal Component Analysis, the most common feature engineering routine in Raman spectroscopy. It is finally applied to a dataset of 223 *in vivo* Raman spectra, revealing key molecular processes that distinguish dense cancer from normal brain tissues.

The first objective of the project was to devise a method that could transform Raman data into a representation that can be submitted to functionally significant statistical analysis. In a functional analysis the focus is put on interpretability of the data in terms that are relevant to the scientists and not the statistician[315]. Here, the Raman spectra are expressed in terms of their functional unit: peaks. This allows not only to simplify the number of variables used to express the spectra, but also to discard unwanted signal. Of course, every simplification risk discarding useful information; here, this is mitigated by identifying the most high-yield regions based on literature research. A famous idiom in statistics states “All models are wrong; some are useful”. In this case, the proposed feature engineering was compared to another “wrong model” of the data: PCA. This analysis revealed that the proposed method is identical to or could possibly surpass PCA in terms of predictive power when used in a predictive analysis of tissue type; this is due to the conservation of more useful, problem-specific information into its representation.

The second objective was to integrate peak-to-peak and peak-to-age interactions as exploratory variables. We chose age as it was the variable that showed the most impact on molecular composition of both normal brain and glioma in the literature. A simple, generalized linear model was used to define a possible relationship between the identified features. While the relatively small dataset size precluded the inclusion of too many features, the feature extraction technique along with the strong regularizing priors made possible to test many hypotheses in a single model. The

representation that resulted is analogous to a graph network, where each peak and clinical attribute is a node and the interactions, the links.

The third objective was to apply the feature engineering pipeline to a dataset of Raman spectra acquired on brain tissue of human patients during surgical resection of glioma. Intraoperative Raman spectra are extremely hard to acquire because of low Raman signal, high tissue autofluorescence, poorly controlled experimental conditions and short acquisition times[47]. The method presented here was designed precisely for this type of data: ones that are acquired in high-uncertainty environment. The features most associated with cancer related to the presence of nucleic acids and several protein compounds, specifically collagen, tryptophan and phenylalanine. Age seemed to interact with features associated to both proteins, lipids and nucleic acids. Importantly, because it relied on Bayesian inference, the analysis could quantify the uncertainty associated with each of its estimates. This allows the scientist to select the level of certainty he is most concerned with; in this case, as this was an exploratory analysis, a high sensitivity and low specificity threshold is better to avoid rejecting potentially interesting features at the cost of detecting non-useful ones.

5.1 Limitations

Several limitations need to be highlighted.

First, the **generalizability** of the results to new data can only be estimated, as their validity was not tested on independent tissue samples. Instead, all available data were used to increase statistical power and detect small effect sizes. The molecular significance of the detected spectral features is based on literature and previous experiments but remains, for some bands, vague. For example, is the prominence of the 1659 cm^{-1} band in cancer a marker of nucleic acids, tryptophan, or Amide I? Moreover, the technique was designed to improve insight gained from Raman data; improvement in predictive performances in classification tasks cannot be extrapolated.

Second, the **peak fitting** algorithm could be improved; 18.8% of all peaks could not be evaluated in high quality spectra. In those cases, a conservative approach to missing values was used, resulting in lower statistical power. In a data-scarce setting where obtaining high-quality observations is so valuable, improving the quality of the fit could be an effective way to maximize data utilization and accelerate experiments.

Third, the **data collection** itself could have introduced biases and could be significantly improved. Data was collected throughout five years by a single surgeon; this individual gained experience in probe manipulation, tissue cleaning before data acquisition, handling of external light sources, and other factors that can affect the shape of the Raman signal, which would be different for a surgeon using the probe for the first time. Brain sample labeling was also performed by a single neuropathologist. These labels were assigned on a qualitative basis, visually estimating the percentage of cancer cells occupying the extracted tissue. A more rigorous approach would be to directly quantify the amount of cancer cells present; this way, a predictive model could be trained to predict the quantity of cancer in a sample, and then emit a diagnosis based on this quantity. This approach to predictive modeling is considered superior to directly estimating the end-result[256]: it increases the amount of information available, exposes uncertainties in outcome (for example, uncertainty is different with 55% cancer infiltration vs. 99% cancer infiltration) and allows more precise and flexible modeling.

5.2 Importance

The project's scientific value resides in four aspects:

- 1) **The importance of the scientific question that it addresses.** While multiple studies have shown that Raman can distinguish normal brain from cancer, this is the first study that focuses on *how* the Raman systems make this distinction. Understanding the inner workings of a new diagnostic tool is essential for acceptance by clinicians and scientists[50].
- 2) **A model designed specifically for *in vivo* application.** Raman data analysis is a long-standing field that has its roots in chemistry and solid physics, where the models of light-material interaction are very tightly controlled and defined. These techniques are now being translated to the chaotic and complex reality of *in vivo* biological systems. Our approach put a specific emphasis on the biological processes generating the Raman signal. This strategy potentially allows to integrate specific information such as clinical variables, molecular pathways, and gene expression. In this fashion, our work is inspired by other data-driven fields such as genomics, proteomics and neuroscience[316]. Raman spectroscopy has improved our knowledge of living tissues, but living tissues can also improve our understanding of Raman spectroscopy.

- 3) **A focus on analytical robustness.** Studies of Raman spectroscopy in human are performed on a restricted number of patients. Nevertheless, measures of model performances, variable importance or other estimations are disclosed as point-values[175]. In the clinical translation phase of a technology, accounting for the additional uncertainty and heterogeneity of effects inherent to the study of human subjects is critical[317]–[319]. Here, the Bayesian framework allows to quantify this uncertainty, both when comparing with other models (e.g. PCA) and when trying to identify the most useful features in a dataset.
- 4) **The quality *and* quantity of the data.** Throughout the development of an imaging device, the signal it acquires can vary significantly. This creates an important challenge when building predictive models; inevitably, early data of lower quality will always be used to train model, which will be tested with the latest version of the technology. Here, even though data was acquired with different systems on a five-year interval, the Raman signal exhibits the same spectral features in spectra with high Raman SNR. This displays the importance of real-time quality-check during data acquisition, an active area of research in our group. By stratifying the spectra by data-quality, we were able to include patients from the entire time-range of the development of the device, for a total of 59 patients. This is the largest number of patients included in a human brain tissue Raman study[225].

CHAPTER 6 FUTURE IMPLICATIONS AND CONCLUSION

The feature engineering method developed during this project, as well as the results that emanated from its application to a dataset of human patients, opens a wide range of potential application and hypotheses.

While the feature engineering pipeline highlighted the potential importance of including interaction terms and clinical variables into predictive models, its capacity to retrieve molecular composition needs to be validated in an environment where the exact molecular composition is known. A good starting point would be to obtain the Raman spectra of cultured cell and vary their molecular composition by altering the genetic expression of specific amino acids and proteins, such as phenylalanine or collagen. This kind of experiment requires expertise in molecular biology, and risks creating a biological model that behaves differently than *in situ* human tissue.

Furthermore, the feature engineering pipeline, by improving our representation of Raman data, could be exploited to enhance performances of predictive machine learning models. As shown in Figure 4.5 : Comparison of reconstruction error and cancer predictive performances between the Raman representation resulting from band fitting and dimensionality reduction following Principal Component Analysis (PCA)., the model can capture a representation of the Raman data that is very different to the one captured by other algorithms (in this case: PCA). In model aggregating, different infant classifiers, with different learning strategies, are trained on the same data, and their prediction are then combined (e.g. through a parent classification algorithm) to emit final predictions for a sample. Presenting different representation to the infant classifier can significantly increase the predictive performances of the parent classifier (through its effect on the model's variance)[320].

Finally, while this work addresses the specific problem of diagnosing glioma during neurosurgeries with Raman spectroscopy, the ideas can theoretically extend to other type of tissues and clinical challenges. Raman spectroscopy is used to detect cancer cells on the skin, prostate, uterus, breast, lung, and blood. It is also used with *ex vivo* samples analyzed by pathologists to accelerate their workflow and enhance their capacity to detect disease-specific biomarkers. The feature engineering pipeline could be applied to Raman signal from these different tissue types (e.g. prostate, skin, blood, *ex vivo* tissues) and different clinical problematics (e.g. chemotherapy responders vs. non-

responders, cancer recurrence *vs.* remission, presence of mutated gene *vs.* wildtype gene). These ideas are currently being explored in various projects by our group.

Raman spectroscopy is a powerful technology that promises to fill a gap between highly specific molecular detection and portable, hand-held tools for real-time surgical guidance during cancer resection. Improving analytical techniques used to model the Raman signal leads to a better understanding of cancer-specific Raman generating processes and clinical confounding variables. In the era of Big data and personalized medicine, technique specific-algorithms can increase the scope of Raman-based systems, enhance their impact on biological research and accelerate their transition into the clinical setting.

BIBLIOGRAPHIE

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA. Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.
- [2] D. Hanahan and R. A. Weinberg, “The Hallmarks of Cancer,” *Cell*, vol. 100, no. 1, pp. 57–70, Jan. 2000, doi: 10.1016/S0092-8674(00)81683-9.
- [3] D. Hanahan and R. A. Weinberg, “Hallmarks of Cancer: The Next Generation,” *Cell*, vol. 144, no. 5, pp. 646–674, Mar. 2011, doi: 10.1016/j.cell.2011.02.013.
- [4] Q. T. Ostrom *et al.*, “CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2010–2014,” *Neuro-Oncol.*, vol. 19, no. Suppl 5, pp. v1–v88, Nov. 2017, doi: 10.1093/neuonc/nox158.
- [5] S. Lapointe, A. Perry, and N. A. Butowski, “Primary brain tumours in adults,” *The Lancet*, vol. 392, no. 10145, pp. 432–446, Aug. 2018, doi: 10.1016/S0140-6736(18)30990-5.
- [6] D. N. Louis *et al.*, “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary,” *Acta Neuropathol. (Berl.)*, vol. 131, no. 6, pp. 803–820, Jun. 2016, doi: 10.1007/s00401-016-1545-1.
- [7] M. M. Kim, A. Parolia, M. P. Dunphy, and S. Veneti, “Non-invasive metabolic imaging of brain tumours in the era of precision medicine,” *Nat. Rev. Clin. Oncol.*, vol. 13, pp. 725–739, Jul. 2016, doi: 10.1038/nrclinonc.2016.108.
- [8] M. Strickland and E. A. Stoll, “Metabolic Reprogramming in Glioma,” *Front. Cell Dev. Biol.*, vol. 5, p. 43, Apr. 2017, doi: 10.3389/fcell.2017.00043.
- [9] M. D. Jenkinson *et al.*, “Intraoperative imaging technology to maximise extent of resection for glioma,” *Cochrane Database Syst. Rev.*, Jan. 2018, doi: 10.1002/14651858.CD012788.pub2.
- [10] N. Kuhar, S. Sil, T. Verma, and S. Umopathy, “Challenges in application of Raman spectroscopy to biology and materials,” *RSC Adv.*, vol. 8, no. 46, pp. 25888–25908, 2018, doi: 10.1039/C8RA04491K.
- [11] R. E. Kass *et al.*, “Computational Neuroscience: Mathematical and Statistical Perspectives,” *Annu. Rev. Stat. Its Appl.*, vol. 5, no. 1, pp. 183–214, 2018, doi: 10.1146/annurev-statistics-041715-033733.
- [12] D. Bzdok, “Classical Statistics and Statistical Learning in Imaging Neuroscience,” *Front. Neurosci.*, vol. 11, 2017, doi: 10.3389/fnins.2017.00543.
- [13] O. Sporns, “Graph theory methods: applications in brain networks,” *Dialogues Clin. Neurosci.*, vol. 20, no. 2, pp. 111–121, Jun. 2018.
- [14] F. V. Farahani, W. Karwowski, and N. R. Lighthall, “Application of Graph Theory for Identifying Connectivity Patterns in Human Brain Networks: A Systematic Review,” *Front. Neurosci.*, vol. 13, 2019, doi: 10.3389/fnins.2019.00585.
- [15] M. A. Beaumont and B. Rannala, “The Bayesian revolution in genetics,” *Nat. Rev. Genet.*, vol. 5, no. 4, pp. 251–261, Apr. 2004, doi: 10.1038/nrg1318.

- [16] M. Stephens and D. J. Balding, “Bayesian statistical methods for genetic association studies,” *Nat. Rev. Genet.*, vol. 10, no. 10, pp. 681–690, Oct. 2009, doi: 10.1038/nrg2615.
- [17] M. W. Libbrecht and W. S. Noble, “Machine learning applications in genetics and genomics,” *Nat. Rev. Genet.*, vol. 16, no. 6, pp. 321–332, Jun. 2015, doi: 10.1038/nrg3920.
- [18] P. Bühlmann, M. Kalisch, and L. Meier, “High-Dimensional Statistics with a View Toward Applications in Biology,” *Annu. Rev. Stat. Its Appl.*, vol. 1, no. 1, pp. 255–278, 2014, doi: 10.1146/annurev-statistics-022513-115545.
- [19] W. Zhang, J. Chien, J. Yong, and R. Kuang, “Network-based machine learning and graph theory algorithms for precision oncology,” *Npj Precis. Oncol.*, vol. 1, no. 1, pp. 1–15, Aug. 2017, doi: 10.1038/s41698-017-0029-7.
- [20] M. Lualdi and M. Fasano, “Statistical analysis of proteomics data: A review on feature selection,” *J. Proteomics*, vol. 198, pp. 18–26, Apr. 2019, doi: 10.1016/j.jprot.2018.12.004.
- [21] A. Schmidt, I. Forne, and A. Imhof, “Bioinformatic analysis of proteomics data,” *BMC Syst. Biol.*, vol. 8, no. Suppl 2, p. S3, Mar. 2014, doi: 10.1186/1752-0509-8-S2-S3.
- [22] E. Cordero, I. Latka, C. Matthäus, I. W. Schie, and J. Popp, “In-vivo Raman spectroscopy: from basics to applications,” *J. Biomed. Opt.*, vol. 23, no. 7, pp. 1–23, 2018.
- [23] O. R. Scepanovic *et al.*, “Multimodal spectroscopy detects features of vulnerable atherosclerotic plaque,” *J. Biomed. Opt.*, vol. 16, no. 1, pp. 1–10, Jan. 2011.
- [24] J. T. Motz *et al.*, “In vivo Raman spectral pathology of human atherosclerosis and vulnerable plaque,” *J. Biomed. Opt.*, vol. 11, no. 2, pp. 1–9, Mar. 2006.
- [25] X. Bi, A. Walsh, A. Mahadevan-Jansen, and A. Herline, “Development of Spectral Markers for the Discrimination of Ulcerative Colitis and Crohn’s Disease Using Raman Spectroscopy,” *Dis. Colon Rectum*, vol. 54, no. 1, 2011, [Online]. Available: https://journals.lww.com/dcrjournal/Fulltext/2011/01000/Development_of_Spectral_Markers_for_the.9.aspx.
- [26] I. J. Pence *et al.*, “Clinical characterization of in vivo inflammatory bowel disease with Raman spectroscopy,” *Biomed. Opt. Express*, vol. 8, no. 2, pp. 524–535, Jan. 2017, doi: 10.1364/BOE.8.000524.
- [27] J. Wang *et al.*, “Fiber-optic Raman spectroscopy for in vivo diagnosis of gastric dysplasia,” *Faraday Discuss.*, vol. 187, no. 0, pp. 377–392, 2016, doi: 10.1039/C5FD00151J.
- [28] J. Wang *et al.*, “Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy improves in vivo diagnosis of esophageal squamous cell carcinoma at endoscopy,” *Sci. Rep.*, vol. 5, p. 12957, Aug. 2015.
- [29] N. D. Magee, J. S. Villaumie, E. T. Marple, M. Ennis, J. S. Elborn, and J. J. McGarvey, “Ex Vivo Diagnosis of Lung Cancer Using a Raman Miniprobe,” *J. Phys. Chem. B*, vol. 113, no. 23, pp. 8137–8141, Jun. 2009, doi: 10.1021/jp900379w.
- [30] H. C. McGregor *et al.*, “Real-time endoscopic Raman spectroscopy for in vivo early lung cancer detection,” *J Biophotonics*, vol. 10, no. 1, pp. 98–110, 2016.
- [31] B. Brożek-Pluska, I. Placek, K. Kurczewski, Z. Morawiec, M. Tazbir, and H. Abramczyk, “Breast cancer diagnostics by Raman spectroscopy,” *Recent Adv. Laser Spectrosc. Laser Technol.*,

vol. 141, no. 3, pp. 145–148, Jun. 2008, doi: 10.1016/j.molliq.2008.02.015.

[32] H. Abramczyk, B. Brozek-Pluska, J. Surmacki, J. Jablonska, and R. Kordek, “The label-free Raman imaging of human breast cancer,” *Complex Liq. Mod. Trends Explor. Underst. Appl. Sel. Pap. Mol. Liq. Present. EMLGJMLG Annu. Meet. 5-9 Sept. 2010*, vol. 164, no. 1, pp. 123–131, Nov. 2011, doi: 10.1016/j.molliq.2011.04.021.

[33] Q. Li, Q. Gao, and G. Zhang, “Classification for breast cancer diagnosis with Raman spectroscopy,” *Biomed. Opt. Express*, vol. 5, no. 7, pp. 2435–2445, Jun. 2014, doi: 10.1364/BOE.5.002435.

[34] A. S. Haka *et al.*, “Diagnosing breast cancer using Raman spectroscopy: prospective analysis,” *J. Biomed. Opt.*, vol. 14, no. 5, pp. 1–8, Sep. 2009.

[35] H. Lui, J. Zhao, D. McLean, and H. Zeng, “Real-time Raman Spectroscopy for *In Vivo* Skin Cancer Diagnosis,” *Cancer Res.*, vol. 72, no. 10, p. 2491, May 2012, doi: 10.1158/0008-5472.CAN-11-4061.

[36] L. Lim *et al.*, “Clinical study of noninvasive *in vivo* melanoma and nonmelanoma skin cancers using multimodal spectral diagnosis,” *J. Biomed. Opt.*, vol. 19, no. 11, pp. 1–12, Nov. 2014.

[37] J. Schleusener *et al.*, “*In vivo* study for the discrimination of cancerous and normal skin using fibre probe-based Raman spectroscopy,” *Exp. Dermatol.*, vol. 24, no. 10, pp. 767–772, Oct. 2015, doi: 10.1111/exd.12768.

[38] B. W. D. de Jong *et al.*, “Discrimination between Nontumor Bladder Tissue and Tumor by Raman Spectroscopy,” *Anal. Chem.*, vol. 78, no. 22, pp. 7761–7769, Nov. 2006, doi: 10.1021/ac061417b.

[39] P. Crow, A. Molckovsky, N. Stone, J. Uff, B. Wilson, and L.-M. WongKeeSong, “Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer,” *Urology*, vol. 65, no. 6, pp. 1126–1130, Jun. 2005, doi: 10.1016/j.urology.2004.12.058.

[40] M. Pinto *et al.*, “Integration of a Raman spectroscopy system to a robotic-assisted surgical system for real-time tissue characterization during radical prostatectomy procedures,” *J. Biomed. Opt.*, vol. 24, no. 2, pp. 1–10, Feb. 2019.

[41] S. Duraipandian, W. Zheng, J. Ng, J. J. H. Low, A. Ilancheran, and Z. Huang, “*In vivo* diagnosis of cervical precancer using Raman spectroscopy and genetic algorithm techniques,” *Analyst*, vol. 136, no. 20, pp. 4328–4336, 2011, doi: 10.1039/C1AN15296C.

[42] J. Mo, W. Zheng, J. J. H. Low, J. Ng, A. Ilancheran, and Z. Huang, “High Wavenumber Raman Spectroscopy for *In Vivo* Detection of Cervical Dysplasia,” *Anal. Chem.*, vol. 81, no. 21, pp. 8908–8915, Nov. 2009, doi: 10.1021/ac9015159.

[43] M. Jermyn *et al.*, “Intraoperative brain cancer detection with Raman spectroscopy in humans,” *Sci. Transl. Med.*, vol. 7, no. 274, pp. 274ra19–274ra19, Feb. 2015, doi: 10.1126/scitranslmed.aaa2384.

[44] M. Jermyn *et al.*, “Highly Accurate Detection of Cancer *In Situ* with Intraoperative, Label-Free, Multimodal Optical Spectroscopy,” *Cancer Res.*, vol. 77, no. 14, pp. 3942–3950, Jul. 2017, doi: 10.1158/0008-5472.CAN-17-0668.

- [45] J. Desroches *et al.*, “Development and first in-human use of a Raman spectroscopy guidance system integrated with a brain biopsy needle,” *J. Biophotonics*, vol. 12, no. 3, p. e201800396, Jan. 2019, doi: 10.1002/jbio.201800396.
- [46] A. Beljebbar, S. Dukic, N. Amharref, and M. Manfait, “Ex vivo and in vivo diagnosis of C6 glioblastoma development by Raman spectroscopy coupled to a microprobe.,” *Anal. Bioanal. Chem.*, vol. 398, no. 1, pp. 477–487, 2010, doi: 10.1007/s00216-010-3910-6.
- [47] C. Beleites and R. Salzer, “Assessing and improving the stability of chemometric models in small sample size situations.,” *Anal. Bioanal. Chem.*, vol. 390, no. 5, pp. 1261–71, 2008, doi: 10.1007/s00216-007-1818-6.
- [48] J. Trevisan, P. P. Angelov, P. L. Carmichael, A. D. Scott, and F. L. Martin, “Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives,” *The Analyst*, vol. 137, no. 14, p. 3202, 2012, doi: 10.1039/c2an16300d.
- [49] C. Beleites, K. Geiger, M. Kirsch, S. B. Sobottka, G. Schackert, and R. Salzer, “Raman spectroscopic grading of astrocytoma tissues: using soft reference information.,” *Anal. Bioanal. Chem.*, vol. 400, no. 9, pp. 2801–16, 2011, doi: 10.1007/s00216-011-4985-4.
- [50] J. P. A. Ioannidis, B. Y. S. Kim, and A. Trounson, “How to design preclinical studies in nanomedicine and cell therapy to maximize the prospects of clinical translation,” *Nat. Biomed. Eng.*, vol. 2, no. 11, pp. 797–809, Nov. 2018, doi: 10.1038/s41551-018-0314-y.
- [51] G. Theophilou, M. Paraskevaïdi, K. M. G. Lima, M. Kyrgiou, P. L. Martin-Hirsch, and F. L. Martin, “Extracting biomarkers of commitment to cancer development: potential role of vibrational spectroscopy in systems biology.,” *Expert Rev. Mol. Diagn.*, vol. 15, no. 5, pp. 693–713, 2015, doi: 10.1586/14737159.2015.1028372.
- [52] Michael C. Dewan *et al.*, “Global neurosurgery: the current capacity and deficit in the provision of essential neurosurgical care. Executive Summary of the Global Neurosurgery Initiative at the Program in Global Surgery and Social Change,” *J. Neurosurg. JNS*, vol. 130, no. 4, pp. 1055–1064, 2018, doi: 10.3171/2017.11.JNS171500.
- [53] W. Stummer *et al.*, “Intraoperative detection of malignant gliomas by 5-aminolevulinic acid-induced porphyrin fluorescence,” *Neurosurgery*, vol. 42, no. 3, pp. 518–526, 1998, doi: 10.1097/00006123-199803000-00017.
- [54] W. Stummer, U. Pichlmeier, T. Meinel, O. D. Wiestler, F. Zanella, and H.-J. Reulen, “Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial,” *Lancet Oncol.*, vol. 7, no. 5, pp. 392–401, May 2006, doi: 10.1016/S1470-2045(06)70665-9.
- [55] P. A. Valdes, V. L. Jacobs, B. C. Wilson, F. Leblond, D. W. Roberts, and K. D. Paulsen, “System and methods for wide-field quantitative fluorescence imaging during neurosurgery,” *Opt. Lett.*, vol. 38, no. 15, p. 2786, 2013, doi: 10.1364/ol.38.002786.
- [56] P. A. Valdés, F. Leblond, V. L. Jacobs, B. C. Wilson, K. D. Paulsen, and D. W. Roberts, “Quantitative, spectrally-resolved intraoperative fluorescence imaging,” *Sci. Rep.*, vol. 2, no. 1, p. 798, Dec. 2012, doi: 10.1038/srep00798.
- [57] L. Marcu *et al.*, “Fluorescence lifetime spectroscopy of glioblastoma multiforme,”

Photochem. Photobiol., vol. 80, no. 1, p. 98, 2004, doi: 10.1562/2003-12-09-ra-023.1.

[58] P. V. Butte, A. N. Mamelak, M. Nuno, S. I. Bannykh, K. L. Black, and L. Marcu, “Fluorescence lifetime spectroscopy for guided therapy of brain tumors,” *NeuroImage*, vol. 54, no. SUPPL. 1, pp. S125–S135, Jan. 2011, doi: 10.1016/j.neuroimage.2010.11.001.

[59] S. R. Kantelhardt *et al.*, “In vivo multiphoton tomography and fluorescence lifetime imaging of human brain tumor tissue,” *J. Neurooncol.*, vol. 127, no. 3, pp. 473–482, May 2016, doi: 10.1007/s11060-016-2062-8.

[60] C. Kut *et al.*, “Detection of human brain cancer infiltration ex vivo and in vivo using quantitative optical coherence tomography,” *Sci. Transl. Med.*, vol. 7, no. 292, pp. 292ra100–292ra100, Jun. 2015, doi: 10.1126/scitranslmed.3010611.

[61] H. Fabelo *et al.*, “An intraoperative visualization system using hyperspectral imaging to aid in brain tumor delineation,” *Sens. Switz.*, vol. 18, no. 2, 2018, doi: 10.3390/s18020430.

[62] K. Wårdell, S. Hemm-Ode, P. Rejmstad, and P. Zsigmond, “High-resolution laser doppler measurements of microcirculation in the deep brain structures: A method for potential vessel tracking,” *Stereotact. Funct. Neurosurg.*, vol. 94, no. 1, pp. 1–9, 2016, doi: 10.1159/000442894.

[63] J. D. Johansson, P. Blomstedt, N. Haj-Hosseini, A. T. Bergenheim, O. Eriksson, and K. Wårdell, “Combined diffuse light reflectance and electrical impedance measurements as a navigation aid in deep brain surgery,” *Stereotact. Funct. Neurosurg.*, vol. 87, no. 2, pp. 105–113, 2009, doi: 10.1159/000202977.

[64] C. a Giller, H. Liu, D. C. German, D. Kashyap, and R. B. Dewey, “A stereotactic near-infrared probe for localization during functional neurosurgical procedures: further experience.,” *J. Neurosurg.*, vol. 110, no. 2, pp. 263–273, 2009, doi: 10.3171/2008.8.JNS08728.

[65] D. DePaoli *et al.*, “Intraoperative fiber optic guidance during chronic electrode implantation in deep brain stimulation neurosurgery: proof of concept in primates,” *J. Neurosurg.*, pp. 1–10, 2019, doi: 10.3171/2019.1.JNS182600.

[66] D. T. DePaoli, N. Lapointe, Y. Messaddeq, M. Parent, and D. C. Côté, “Intact primate brain tissue identification using a completely fibered coherent Raman spectroscopy system,” *Neurophotonics*, vol. 5, no. 03, p. 1, 2018, doi: 10.1117/1.NPh.5.3.035005.

[67] J. Pichette *et al.*, “Intraoperative video-rate hemodynamic response assessment in human cortex using snapshot hyperspectral optical imaging,” *Neurophotonics*, vol. 3, no. 4, p. 045003, 2016, doi: 10.1117/1.nph.3.4.045003.

[68] H. Ramakonar *et al.*, “Intraoperative detection of blood vessels with an imaging needle during neurosurgery in humans,” *Sci. Adv.*, vol. 4, no. 12, p. eaav4992, 2018, doi: 10.1126/sciadv.aav4992.

[69] J. Desroches *et al.*, “A new method using Raman spectroscopy for in vivo targeted brain cancer tissue biopsy,” *Sci. Rep.*, vol. 8, no. 1, p. 1792, Dec. 2018, doi: 10.1038/s41598-018-20233-3.

[70] T. K. Owonikoko *et al.*, “Current approaches to the treatment of metastatic brain tumours,” *Nat. Rev. Clin. Oncol.*, vol. 11, p. 203, Feb. 2014.

[71] I. J. Gerard, M. Kersten-Oertel, K. Petrecca, D. Sirhan, J. A. Hall, and D. L. Collins, “Brain

shift in neuronavigation of brain tumors: A review,” *Med. Image Anal.*, vol. 35, pp. 403–420, Jan. 2017, doi: 10.1016/j.media.2016.08.007.

[72] K. M. Brindle, J. L. Izquierdo-García, D. Y. Lewis, R. J. Mair, and A. J. Wright, “Brain Tumor Imaging,” *J. Clin. Oncol.*, vol. 35, no. 21, pp. 2432–2438, Jun. 2017, doi: 10.1200/JCO.2017.72.7636.

[73] T. Reithmeier *et al.*, “Intraindividual comparison of histopathological diagnosis obtained by stereotactic serial biopsy to open surgical resection specimen in patients with intracranial tumours,” *Clin. Neurol. Neurosurg.*, vol. 115, no. 10, pp. 1955–1960, Oct. 2013, doi: 10.1016/j.clineuro.2013.05.019.

[74] M. S. Okun *et al.*, “Management of referred deep brain stimulation failures,” *Arch. Neurol.*, vol. 62, no. 8, pp. 1250–1255, 2005, doi: 10.1001/archneur.62.8.noc40425.

[75] D. W. Shipp, F. Sinjab, and I. Notinger, “Raman spectroscopy: techniques and applications in the life sciences,” *Adv. Opt. Photonics*, vol. 9, no. 2, pp. 315–428, Jun. 2017, doi: 10.1364/AOP.9.000315.

[76] A. Alfonso-García, R. Mittal, E. S. Lee, and E. O. Potma, “Biological imaging with coherent Raman scattering microscopy: a tutorial,” *J. Biomed. Opt.*, vol. 19, no. 7, p. 71407, 2014, doi: 10.1117/1.JBO.19.7.071407.

[77] W. Min, C. W. Freudiger, S. Lu, and X. S. Xie, “Coherent nonlinear optical imaging: Beyond fluorescence microscopy,” *Annu. Rev. Phys. Chem.*, vol. 62, no. 1, pp. 507–530, 2011, doi: 10.1146/annurev.physchem.012809.103512.

[78] D. Zhang, P. Wang, M. N. Slipchenko, and J. X. Cheng, “Fast vibrational imaging of single cells and tissues by stimulated raman scattering microscopy,” *Acc. Chem. Res.*, vol. 47, no. 8, pp. 2282–2290, 2014, doi: 10.1021/ar400331q.

[79] R. L. McCreery, *Raman Spectroscopy for Chemical Analysis*. Wiley, 2005.

[80] R. S. Das and Y. K. Agrawal, “Raman spectroscopy: Recent advancements, techniques and applications,” *Vib. Spectrosc.*, vol. 57, no. 2, pp. 163–176, Nov. 2011, doi: 10.1016/j.vibspec.2011.08.003.

[81] C. L. Evans and X. S. Xie, “Coherent anti-stokes raman scattering microscopy: Chemical imaging for biology and medicine,” *Annu. Rev. Anal. Chem.*, vol. 1, no. 1, pp. 883–909, 2008, doi: 10.1146/annurev.anchem.1.031207.112754.

[82] A. Zumbusch, G. R. Holtom, and X. S. Xie, “Three-dimensional vibrational imaging by coherent anti-Stokes Raman scattering,” *Phys. Rev. Lett.*, vol. 82, no. 20, pp. 4142–4145, 1999, doi: 10.1103/PhysRevLett.82.4142.

[83] J. X. Cheng, Y. K. Jia, G. Zheng, and X. S. Xie, “Laser-scanning coherent anti-Stokes Raman scattering microscopy and applications to cell biology,” *Biophys. J.*, vol. 83, no. 1, pp. 502–509, 2002, doi: 10.1016/S0006-3495(02)75186-2.

[84] E. O. Potma, W. P. de Boeij, P. J. M. van Haastert, and D. A. Wiersma, “Real-time visualization of intracellular hydrodynamics in single living cells,” *Proc. Natl. Acad. Sci.*, vol. 98, no. 4, pp. 1577–1582, 2001, doi: 10.1073/pnas.98.4.1577.

[85] X. Nan, J.-X. Cheng, and X. S. Xie, “Vibrational imaging of lipid droplets in live fibroblast

cells with coherent anti-Stokes Raman scattering microscopy,” *J. Lipid Res.*, vol. 44, no. 11, pp. 2202–2208, 2003, doi: 10.1194/jlr.d300022-jlr200.

[86] E. O. Potma and X. S. Xie, “Detection of single lipid bilayers with coherent anti-Stokes Raman scattering (CARS) microscopy,” *J. Raman Spectrosc.*, vol. 34, no. 9, pp. 642–650, 2003, doi: 10.1002/jrs.1045.

[87] H. Wang, Y. Fu, P. Zickmund, R. Shi, and J.-X. Cheng, “Coherent anti-stokes Raman scattering imaging of axonal myelin in live spinal tissues.,” *Biophys. J.*, vol. 89, no. 1, pp. 581–91, 2005, doi: 10.1529/biophysj.105.061911.

[88] C. L. Evans, E. O. Potma, M. Puoris’haag, D. Côté, C. P. Lin, and X. S. Xie, “Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 46, pp. 16807–16812, 2005, doi: 10.1073/pnas.0508282102.

[89] T. B. Huff and J.-X. Cheng, “In vivo coherent anti-Stokes Raman scattering imaging of sciatic nerve tissue,” *J. Microsc.*, vol. 225, no. 2, pp. 175–182, 2007, doi: 10.1111/j.1365-2818.2007.01729.x.

[90] Y. Fu, T. B. Huff, H.-W. Wang, H. Wang, and J.-X. Cheng, “Ex vivo and in vivo imaging of myelin fibers in mouse brain by coherent anti-Stokes Raman scattering microscopy.,” *Opt. Express*, vol. 16, no. 24, p. 19396, 2008.

[91] E. Bélanger, S. Bégin, S. Laffray, Y. De Koninck, R. Vallée, and D. Côté, “Quantitative myelin imaging with coherent anti-Stokes Raman scattering microscopy: alleviating the excitation polarization dependence with circularly polarized laser beams,” *Opt. Express*, vol. 17, no. 21, p. 18419, 2009, doi: 10.1364/oe.17.018419.

[92] S. Bégin, B. Burgoyne, V. Mercier, A. Villeneuve, R. Vallée, and D. Côté, “Coherent anti-Stokes Raman scattering hyperspectral tissue imaging with a wavelength-swept system,” *Biomed. Opt. Express*, vol. 2, no. 5, p. 1296, 2011, doi: 10.1364/BOE.2.001296.

[93] E. Bélanger, J. Crépeau, S. Laffray, R. Vallée, Y. De Koninck, and D. Côté, “Live animal myelin histomorphometry of the spinal cord with video-rate multimodal nonlinear microendoscopy,” *J. Biomed. Opt.*, vol. 17, no. 2, p. 021107, 2012, doi: 10.1117/1.jbo.17.2.021107.

[94] R. Turcotte, D. J. Rutledge, E. Bélanger, D. Dill, W. B. Macklin, and D. C. Côté, “Intravital assessment of myelin molecular order with polarimetric multiphoton microscopy,” *Sci. Rep.*, vol. 6, pp. 1–8, 2016, doi: 10.1038/srep31685.

[95] C. W. Freudiger *et al.*, “Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy.,” *Science*, vol. 322, no. 5909, pp. 1857–61, 2008, doi: 10.1126/science.1165758.

[96] F. K. Lu *et al.*, “Label-free DNA imaging in vivo with stimulated Raman scattering microscopy,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 43, p. E5902, 2015, doi: 10.1073/pnas.1519382112.

[97] L. Wei *et al.*, “Imaging complex protein metabolism in live organisms by stimulated Raman scattering microscopy with isotope labeling.,” *ACS Chem. Biol.*, vol. 10, no. 3, pp. 901–908, 2015, doi: 10.1021/cb500787b.

- [98] D. Fu, W. Yang, and X. S. Xie, "Label-free imaging of neurotransmitter acetylcholine at neuromuscular junctions with stimulated Raman scattering," *J. Am. Chem. Soc.*, vol. 139, no. 2, pp. 583–586, 2017, doi: 10.1021/jacs.6b10727.
- [99] M. Ji *et al.*, "Rapid, Label-Free Detection of Brain Tumors with Stimulated Raman Scattering Microscopy," *Sci. Transl. Med.*, vol. 5, no. 201, pp. 201ra119-201ra119, Sep. 2013, doi: 10.1126/scitranslmed.3005954.
- [100] M. Ji *et al.*, "Detection of human brain tumor infiltration with quantitative stimulated Raman scattering microscopy," *Sci. Transl. Med.*, vol. 7, no. 309, p. 309ra163, 2015, doi: 10.1126/scitranslmed.aab0195.
- [101] D. A. Orringer *et al.*, "Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy," *Nat. Biomed. Eng.*, vol. 1, no. 2, p. 0027, Feb. 2017, doi: 10.1038/s41551-016-0027.
- [102] L. F. Santos, R. Wolthuis, S. Koljenović, R. M. Almeida, and G. J. Puppels, "Fiber-Optic Probes for in Vivo Raman Spectroscopy in the High-Wavenumber Region," *Anal. Chem.*, vol. 77, no. 20, pp. 6747–6752, Oct. 2005, doi: 10.1021/ac0505730.
- [103] S. Koljenovic *et al.*, "Tissue characterization using high wave number Raman spectroscopy," *J. Biomed. Opt.*, vol. 10, no. 3, p. 031116, 2005, doi: 10.1117/1.1922307.
- [104] S. Koljenović *et al.*, "Raman Spectroscopic Characterization of Porcine Brain Tissue Using a Single Fiber-Optic Probe," *Anal. Chem.*, vol. 79, no. 2, pp. 557–564, Jan. 2007, doi: 10.1021/ac0616512.
- [105] J. Desroches *et al.*, "Characterization of a Raman spectroscopy probe system for intraoperative brain tissue classification," *Biomed. Opt. Express*, vol. 6, no. 7, p. 2380, Jul. 2015, doi: 10.1364/BOE.6.002380.
- [106] M. Jermyn *et al.*, "Neural networks improve brain cancer detection with Raman spectroscopy in the presence of operating room light artifacts," *J. Biomed. Opt.*, vol. 21, no. 9, p. 094002, Sep. 2016, doi: 10.1117/1.JBO.21.9.094002.
- [107] M. Jermyn *et al.*, "Raman spectroscopy detects distant invasive brain cancer cells centimeters beyond MRI capability in humans," *Biomed. Opt. Express*, vol. 7, no. 12, pp. 5129–5137, 2016, doi: 10.1364/BOE.7.005129.
- [108] O. A. C. Stevens, J. Hutchings, W. Gray, R. L. Vincent, and J. C. Day, "Miniature standoff Raman probe for neurosurgical applications," *J. Biomed. Opt.*, vol. 21, no. 8, p. 87002, 2016, doi: 10.1117/1.JBO.21.8.087002.
- [109] B. Burgoyne, N. Godbout, and S. Lacroix, "Nonlinear pulse propagation in optical fibers using second order moments," *Opt. Express*, vol. 15, no. 16, p. 10075, 2007, doi: 10.1364/oe.15.010075.
- [110] M. T. Cicerone and C. H. Camp, "Potential roles for spectroscopic coherent raman imaging for histopathology and biomedicine," in *Neurophotronics and biomedical spectroscopy*, Elsevier Ltd, 2019, pp. 547–570.
- [111] Z. Li, W. Minzhu, Z. Jian, and Z. Ju, "Application of infrared spectroscopy in biomedical polymer materials," in *Macro to nano spectroscopy*, J. Uddin, Ed. Intechopen, 2012.

- [112] W. Wang, M. Short, I. T. Tai, and H. Zeng, “Disposable sheath that facilitates endoscopic Raman spectroscopy,” *J. Biomed. Opt.*, vol. 21, no. 2, p. 025001, 2016, doi: 10.1117/1.jbo.21.2.025001.
- [113] N. Haj-Hosseini, J. C. O. Richter, P. Milos, M. Hallbeck, and K. Wårdell, “5-ALA fluorescence and laser Doppler flowmetry for guidance in a stereotactic brain tumor biopsy,” *Biomed. Opt. Express*, vol. 9, no. 5, p. 2284, 2018, doi: 10.1364/boe.9.002284.
- [114] F. Picot *et al.*, “Interstitial imaging with multiple diffusive reflectance spectroscopy projections for in vivo blood vessels detection during brain needle biopsy procedures,” *Neurophotonics*, vol. 6, no. 02, p. 1, 2019, doi: 10.1117/1.nph.6.2.025003.
- [115] N. Bergner, C. Krafft, K. D. Geiger, M. Kirsch, G. Schackert, and J. Popp, “Unsupervised unmixing of Raman microspectroscopic images for morphochemical analysis of non-dried brain tumor specimens,” *Anal. Bioanal. Chem.*, vol. 403, no. 3, pp. 719–25, 2012, doi: 10.1007/s00216-012-5858-1.
- [116] N. Bergner *et al.*, “Hyperspectral unmixing of Raman micro-images for assessment of morphological and chemical parameters in non-dried brain tumor specimens,” *Anal. Bioanal. Chem.*, vol. 405, no. 27, pp. 8719–28, 2013, doi: 10.1007/s00216-013-7257-7.
- [117] R. Kast *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.*, vol. 125, no. 2, pp. 287–295, Nov. 2015, doi: 10.1007/s11060-015-1929-4.
- [118] H. Abramczyk and A. Imiela, “The biochemical, nanomechanical and chemometric signatures of brain cancer,” *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 188, no. cuo, 9602533, pp. 8–19, 2018, doi: 10.1016/j.saa.2017.06.037.
- [119] C. L. Evans, X. Xu, S. Kesari, X. S. Xie, S. T. C. Wong, and G. S. Young, “Chemically-selective imaging of brain structures with CARS microscopy,” *Opt. Express*, vol. 15, no. 19, pp. 12076–12087, Sep. 2007, doi: 10.1364/OE.15.012076.
- [120] O. Uckermann *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*, vol. 9, no. 9, p. e107115, 2014, doi: 10.1371/journal.pone.0107115.
- [121] B. F. M. Romeike *et al.*, “Coherent anti-Stokes Raman scattering and two photon excited fluorescence for neurosurgery,” *Clin. Neurol. Neurosurg.*, vol. 131, no. df4, 7502039, pp. 42–6, 2015, doi: 10.1016/j.clineuro.2015.01.022.
- [122] R. Galli *et al.*, “Assessing the efficacy of coherent anti-Stokes Raman scattering microscopy for the detection of infiltrating glioblastoma in fresh brain samples,” *J. Biophotonics*, vol. 10, no. 3, pp. 404–414, 2017, doi: 10.1002/jbio.201500323.
- [123] R. Galli *et al.*, “Identification of distinctive features in human intracranial tumors by label-free nonlinear multimodal microscopy,” *J. Biophotonics*, no. June, p. e201800465, 2019, doi: 10.1002/jbio.201800465.
- [124] A. H. Fischer, K. A. Jacobson, J. Rose, and R. Zeller, “Hematoxylin and eosin staining of tissue and cell sections,” *Cold Spring Harb. Protoc.*, vol. 3, no. 5, pp. 4986–4988, 2008, doi: 10.1101/pdb.prot4986.
- [125] C. W. Freudiger *et al.*, “Multicolored stain-free histopathology with coherent Raman

imaging.,” *Lab. Investig. J. Tech. Methods Pathol.*, vol. 92, no. 10, pp. 1492–1502, 2012, doi: 10.1038/labinvest.2012.109.

[126] B. G. Saar, C. W. Freudiger, J. Reichman, C. M. Stanley, G. R. Holtom, and X. S. Xie, “Video-rate molecular imaging in vivo with stimulated Raman scattering,” *Science*, vol. 330, no. 6009, pp. 1368–1370, 2010, doi: 10.1126/science.1197236.

[127] F.-K. Lu *et al.*, “Label-Free Neurosurgical Pathology with Stimulated Raman Imaging.,” *Cancer Res.*, vol. 76, no. 12, pp. 3451–3462, 2016, doi: 10.1158/0008-5472.CAN-16-0270.

[128] F.-K. Lu, S. Santagata, X. S. Xie, A. J. Golby, and N. Y. R. Agar, “Label-free neurosurgical pathology with stimulated raman imaging: A dataset,” 2016, doi: 10.7910/DVN/EZW4EK.

[129] C. W. Freudiger, W. Yang, G. R. Holtom, N. Peyghambarian, X. S. Xie, and K. Q. Kieu, “Stimulated Raman scattering microscopy with a robust fibre laser source,” *Nat. Photonics*, vol. 8, no. 2, pp. 153–159, 2014, doi: 10.1038/nphoton.2013.360.

[130] T. C. Hollon *et al.*, “Rapid Intraoperative Diagnosis of Pediatric Brain Tumors Using Stimulated Raman Histology.,” *Cancer Res.*, vol. 78, no. 1, pp. 278–289, 2018, doi: 10.1158/0008-5472.CAN-17-1974.

[131] K. Bae, W. Zheng, K. Lin, S. W. Lim, Y. K. Chong, and C. Tang, “Epi-detected hyperspectral stimulated raman scattering microscopy for label-free molecular subtyping of glioblastomas,” *Anal. Chem.*, vol. 90, no. 17, pp. 10249–10255, 2018, doi: 10.1021/acs.analchem.8b01677.

[132] K. St-Arnaud *et al.*, “Development and characterization of a handheld hyperspectral Raman imaging probe system for molecular characterization of tissue on mesoscopic scales,” *Med. Phys.*, vol. 45, no. 1, pp. 328–339, 2018, doi: 10.1002/mp.12657.

[133] K. St-Arnaud *et al.*, “Wide-field spontaneous Raman spectroscopy imaging system for biological tissue interrogation,” *Opt. Lett.*, vol. 41, no. 20, p. 4692, 2016, doi: 10.1364/ol.41.004692.

[134] F. Légaré, C. L. Evans, F. Ganikhanov, and X. S. Xie, “Towards CARS endoscopy,” *Opt Express*, vol. 14, no. 10, pp. 4427–4432, May 2006, doi: 10.1364/OE.14.004427.

[135] M. Balu, G. Liu, Z. Chen, B. J. Tromberg, and E. O. Potma, “Fiber delivered probe for efficient CARS imaging of tissues.,” *Opt. Express*, vol. 18, no. 3, pp. 2380–2388, 2010, doi: 10.1364/OE.18.002380.

[136] P. Deladurantaye *et al.*, “Advances in engineering of high contrast CARS imaging endoscopes,” *Opt. Express*, vol. 22, no. 21, p. 25053, 2014, doi: 10.1364/oe.22.025053.

[137] H. Wang, T. B. Huff, and J. X. Cheng, “Coherent anti-Stokes Raman scattering imaging with a laser source delivered by a photonic crystal fiber,” *Opt Lett*, vol. 31, no. 10, pp. 1417–1419, 2006.

[138] S. Brustlein *et al.*, “Double-clad hollow core photonic crystal fiber for coherent Raman endoscope.,” *Opt. Express*, vol. 19, no. 13, pp. 12562–12568, 2011, doi: 10.1364/OE.19.012562.

[139] A. Lombardini, E. R. Andresen, A. Kudlinski, I. Rimke, and H. Rigneault, “Origin and suppression of parasitic signals in Kagomé lattice hollow core fibers used for SRS microscopy and endoscopy,” *Opt. Lett.*, vol. 42, no. 9, p. 1824, 2017, doi: 10.1364/ol.42.001824.

- [140] A. Lukic *et al.*, “Endoscopic fiber probe for nonlinear spectroscopic imaging,” *Optica*, vol. 4, no. 5, p. 496, 2017, doi: 10.1364/optica.4.000496.
- [141] A. Lombardini *et al.*, “High-resolution multimodal flexible coherent Raman endoscope article,” *Light Sci. Appl.*, vol. 7, no. 1, 2018, doi: 10.1038/s41377-018-0003-3.
- [142] P. Zirak *et al.*, “Invited Article: A rigid coherent anti-Stokes Raman scattering endoscope with high resolution and a large field of view,” *APL Photonics*, vol. 3, no. 9, 2018, doi: 10.1063/1.5027182.
- [143] B. G. Saar, R. S. Johnston, C. W. Freudiger, X. S. Xie, and E. J. Seibel, “Coherent Raman scanning fiber endoscopy,” *Opt. Lett.*, vol. 36, no. 13, pp. 2396–2398, 2011, doi: 10.1364/OL.36.002396.
- [144] C.-S. Liao *et al.*, “Stimulated Raman spectroscopic imaging by microsecond delay-line tuning,” *Optica*, vol. 3, no. 12, p. 1377, 2016, doi: 10.1364/OPTICA.3.001377.
- [145] C. S. Liao *et al.*, “In vivo and in situ spectroscopic imaging by a handheld stimulated raman scattering microscope,” *ACS Photonics*, vol. 5, no. 3, pp. 947–954, 2018, doi: 10.1021/acsp Photonics.7b01214.
- [146] J. X. Cheng and X. S. Xie, “Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine,” *Science*, vol. 350, no. 6264, 2015, doi: 10.1126/science.aaa8870.
- [147] T. Gottschall, T. Meyer, M. Schmitt, J. Popp, J. Limpert, and A. Tünnermann, “Advances in laser concepts for multiplex, coherent Raman scattering micro-spectroscopy and imaging,” *TrAC - Trends Anal. Chem.*, vol. 102, pp. 103–109, 2018, doi: 10.1016/j.trac.2018.01.010.
- [148] Laser Institute of America, “ANSI Z136.1-2014,” in *SMPTE journal*, vol. 90, 2014.
- [149] ICNIRP, “Icnirp guidelines on limits of exposure to laser radiation of wavelengths between 180 nm and 1,000 μm ,” *Health Phys.*, vol. 105, no. 3, pp. 271–295, 2013, doi: 10.1097/HP.0b013e3182983fd4.
- [150] S. L. Jacques, “Optical properties of biological tissues: a review,” *Phys. Med. Biol.*, vol. 58, no. 11, p. R37, 2013.
- [151] M. Nourhashemi, M. Mahmoudzadeh, and F. Wallois, “Thermal impact of near-infrared laser in advanced noninvasive optical brain imaging,” *Neurophotonics*, vol. 3, no. 1, p. 015001, 2016, doi: 10.1117/1.NPh.3.1.015001.
- [152] H. Wang *et al.*, “Brain temperature and its fundamental properties: A review for clinical neuroscientists,” *Front. Neurosci.*, vol. 8, no. SEP, pp. 1–17, 2014, doi: 10.3389/fnins.2014.00307.
- [153] E. A. Kiyatkin, P. L. Brown, and R. A. Wise, “Brain temperature fluctuation: A reflection of functional neural activation,” *Eur. J. Neurosci.*, vol. 16, no. 1, pp. 164–168, 2002, doi: 10.1046/j.1460-9568.2002.02066.x.
- [154] S. Thomsen, “Pathologic analysis of photothermal and photomechanical effects of Laser–Tissue interactions,” *Photochem. Photobiol.*, vol. 53, no. 6, pp. 825–835, 1991, doi: 10.1111/j.1751-1097.1991.tb09897.x.
- [155] A. Hopt and E. Neher, “Highly nonlinear photodamage in two-photon fluorescence microscopy,” *Biophys. J.*, vol. 80, no. April, pp. 1–8, 2001.

- [156] H. J. Koester, D. Baur, R. Uhl, and S. W. Hell, "Ca²⁺ fluorescence imaging with pico- and femtosecond two-photon excitation: Signal and photodamage," *Biophys. J.*, vol. 77, no. 4, pp. 2226–2236, 1999, doi: 10.1016/S0006-3495(99)77063-3.
- [157] A. Vogel, J. Noack, G. Hüttman, and G. Paltauf, "Mechanisms of femtosecond laser nanosurgery of cells and tissues," *Appl. Phys. B Lasers Opt.*, vol. 81, no. 8, pp. 1015–1047, 2005, doi: 10.1007/s00340-005-2036-6.
- [158] K. Podgorski and G. Ranganathan, "Brain heating induced by near-infrared lasers during multiphoton microscopy," *J. Neurophysiol.*, vol. 116, no. 3, pp. 1012–1023, 2016, doi: 10.1152/jn.00275.2016.
- [159] A. C. S. Talari, Z. Movasaghi, S. Rehman, and I. ur Rehman, "Raman Spectroscopy of Biological Tissues," *Appl. Spectrosc. Rev.*, vol. 50, no. 1, pp. 46–111, Jan. 2015, doi: 10.1080/05704928.2014.923902.
- [160] H. J. Butler *et al.*, "Using Raman spectroscopy to characterize biological materials," *Nat. Protoc.*, vol. 11, no. 4, pp. 664–687, Mar. 2016, doi: 10.1038/nprot.2016.036.
- [161] M. Brusatori, G. Auner, T. Noh, L. Scarpace, B. Broadbent, and S. N. Kalkanis, "Intraoperative Raman Spectroscopy," *Neurosurg. Clin. N. Am.*, vol. 28, no. 4, pp. 633–652, Oct. 2017, doi: 10.1016/j.nec.2017.05.014.
- [162] A. C. Croce and G. Bottiroli, "Autofluorescence spectroscopy and imaging: a tool for biomedical research and diagnosis," *Eur. J. Histochem.*, vol. 58, no. 4, p. 18, 2014.
- [163] S. Yang, B. Li, M. N. Slipchenko, A. Akkus, N. G. Singer, and O. Akkus, "Laser Wavelength Dependence of Background Fluorescence in Raman Spectroscopic Analysis of Synovial Fluid from Symptomatic Joints," *J. Raman Spectrosc.*, vol. 44, no. 8, p. 16, 2014.
- [164] H. J. Byrne, P. Knief, M. E. Keating, and F. Bonnier, "Spectral pre and post processing for infrared and Raman spectroscopy of biological tissues and cells," *Chem. Soc. Rev.*, vol. 45, no. 7, pp. 1865–1878, Mar. 2016, doi: 10.1039/C5CS00440C.
- [165] J. Zhao, H. Lui, D. I. McLean, and H. Zeng, "Automated Autofluorescence Background Subtraction Algorithm for Biomedical Raman Spectroscopy," *Appl. Spectrosc.*, vol. 61, no. 11, pp. 1225–1232, Nov. 2007, doi: 10.1366/000370207782597003.
- [166] S. He *et al.*, "Baseline correction for Raman spectra using an improved asymmetric least squares method," *Anal. Methods*, vol. 6, no. 12, p. 6, 2014.
- [167] S. Guo, T. Bocklitz, and J. Popp, "Optimization of Raman-spectrum baseline correction in biological application," *The Analyst*, vol. 141, no. 8, pp. 2396–2404, 2016, doi: 10.1039/C6AN00041J.
- [168] Y. Cai, C. Yang, D. Xu, and W. Gui, "Baseline correction for Raman spectra using penalized spline smoothing based on vector transformation," *Anal. Methods*, vol. 10, no. 28, pp. 3525–3533, Jul. 2018, doi: 10.1039/C8AY00914G.
- [169] C. Krafft *et al.*, "Advances in optical biopsy--correlation of malignancy and cell density of primary brain tumors using Raman microspectroscopic imaging.," *The Analyst*, vol. 137, no. 23, pp. 5533–7, 2012, doi: 10.1039/c2an36083g.
- [170] M. J. Pelletier, "Quantitative Analysis Using Raman Spectrometry," *Appl. Spectrosc.*, vol.

57, no. 1, pp. 20A-42A, Jan. 2003.

[171] Y. Bengio, A. Courville, and P. Vincent, “Representation Learning: A Review and New Perspectives,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013, doi: 10.1109/TPAMI.2013.50.

[172] R. P. Aguiar, L. Silveira, E. T. Falcão, M. T. T. Pacheco, R. A. Zângaro, and C. A. Pasqualucci, “Discriminating Neoplastic and Normal Brain Tissues *in Vitro* Through Raman Spectroscopy: A Principal Components Analysis Classification Model,” *Photomed. Laser Surg.*, vol. 31, no. 12, pp. 595–604, Dec. 2013, doi: 10.1089/pho.2012.3460.

[173] E. Ostertag, M. Stefanakis, K. Rebner, and R. W. Kessler, “Elastic and inelastic light scattering spectroscopy and its possible use for label-free brain tumor typing,” *Anal. Bioanal. Chem.*, vol. 409, no. 28, pp. 6613–6623, 2017, doi: 10.1007/s00216-017-0614-1.

[174] E. Kaur *et al.*, “Unique spectral markers discern recurrent Glioblastoma cells from heterogeneous parent population,” *Sci. Rep.*, vol. 6, no. 101563288, p. 26538, 2016, doi: 10.1038/srep26538.

[175] Y. Zhou *et al.*, “Human brain cancer studied by resonance Raman spectroscopy,” *J. Biomed. Opt.*, vol. 17, no. 11, p. 116021, 2012, doi: 10.1117/1.JBO.17.11.116021.

[176] N. Amharref *et al.*, “Discriminating healthy from tumor and necrosis tissue in rat brain tissue samples by Raman spectral imaging,” *Biochim. Biophys. Acta*, vol. 1768, no. 10, pp. 2605–15, 2007.

[177] S. Koljenovic, L.-P. Choo-Smith, T. C. Bakker Schut, J. M. Kros, H. J. van den Berge, and G. J. Puppels, “Discriminating vital tumor from necrotic tissue in human glioblastoma tissue samples by Raman spectroscopy,” *Lab. Investig. J. Tech. Methods Pathol.*, vol. 82, no. 10, pp. 1265–77, 2002.

[178] L. M. Fullwood *et al.*, “Investigating the use of Raman and immersion Raman spectroscopy for spectral histopathology of metastatic brain cancer and primary sites of origin,” *Anal. Methods*, vol. 6, no. 12, pp. 3948–3961, 2014, doi: 10.1039/c3ay42190b.

[179] N. Bergner *et al.*, “Identification of primary tumors of brain metastases by Raman imaging and support vector machines,” *Chemom. Intell. Lab. Syst.*, vol. 117, pp. 224–32, Aug. 2012, doi: 10.1016/j.chemolab.2012.02.008.

[180] K. Gajjar *et al.*, “Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis,” *Anal Methods*, vol. 5, no. 1, pp. 89–102, 2013, doi: 10.1039/C2AY25544H.

[181] S. Wold, A. Ruhe, H. Wold, and I. Dunn W., “The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses,” *SIAM J. Sci. Stat. Comput.*, vol. 5, no. 3, pp. 735–743, Sep. 1984, doi: 10.1137/0905052.

[182] M. Robnik-Šikonja and I. Kononenko, “Theoretical and Empirical Analysis of Relief and RRelief,” *Mach. Learn.*, vol. 53, no. 1, pp. 23–69, Oct. 2003, doi: 10.1023/A:1025667309714.

[183] L. Yu and H. Liu, “Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution,” *Proc. Twent. Int. Conf. Mach. Learn.*, vol. 2, p. 8, 2003.

[184] N. Hoque, D. K. Bhattacharyya, and J. K. Kalita, “MIFS-ND: A mutual information-based

feature selection method,” *Expert Syst. Appl.*, vol. 41, no. 14, pp. 6371–6385, Oct. 2014, doi: 10.1016/j.eswa.2014.04.019.

[185] M. Radovic, M. Ghalwash, N. Filipovic, and Z. Obradovic, “Minimum redundancy maximum relevance feature selection approach for temporal gene expression data,” *BMC Bioinformatics*, vol. 18, no. 1, p. 9, Jan. 2017, doi: 10.1186/s12859-016-1423-9.

[186] R. Kohavi and G. H. John, “Wrappers for feature subset selection,” *Relevance*, vol. 97, no. 1, pp. 273–324, Dec. 1997, doi: 10.1016/S0004-3702(97)00043-X.

[187] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *J. R. Stat. Soc. Ser. B Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.

[188] R. Stables *et al.*, “Feature driven classification of Raman spectra for real-time spectral brain tumour diagnosis using sound.,” *The Analyst*, vol. 142, no. 1, pp. 98–109, 2016.

[189] K. Tanahashi *et al.*, “Assessment of tumor cells in a mouse model of diffuse infiltrative glioma by Raman spectroscopy.,” *BioMed Res. Int.*, vol. 2014, no. 101600173, p. 860241, 2014, doi: 10.1155/2014/860241.

[190] A. W. Auner, R. E. Kast, R. Rabah, J. M. Poulik, and M. D. Klein, “Conclusions and data analysis: a 6-year study of Raman spectroscopy of solid tumors at a major pediatric institute.,” *Pediatr. Surg. Int.*, vol. 29, no. 2, pp. 129–40, 2013, doi: 10.1007/s00383-012-3211-6.

[191] O. Uckermann *et al.*, “IDH1 mutation in human glioma induces chemical alterations that are amenable to optical Raman spectroscopy,” *J. Neurooncol.*, vol. 139, no. 2, pp. 261–268, Sep. 2018, doi: 10.1007/s11060-018-2883-8.

[192] S. N. Kalkanis *et al.*, “Raman spectroscopy to distinguish grey matter, necrosis, and glioblastoma multiforme in frozen tissue sections,” *J. Neurooncol.*, vol. 116, no. 3, pp. 477–485, Feb. 2014, doi: 10.1007/s11060-013-1326-9.

[193] A. Jović, K. Brkić, and N. Bogunović, “A review of feature selection methods with applications,” in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, May 2015, pp. 1200–1205, doi: 10.1109/MIPRO.2015.7160458.

[194] Y. Saeys, I. Inza, and P. Larrañaga, “A review of feature selection techniques in bioinformatics,” *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, Aug. 2007, doi: 10.1093/bioinformatics/btm344.

[195] É. Lemoine *et al.*, “Feature engineering applied to intraoperative in vivo Raman spectroscopy sheds light on molecular processes in brain cancer: a retrospective study of 65 patients,” *Analyst*, vol. 144, no. 22, pp. 6517–6532, 2019, doi: 10.1039/C9AN01144G.

[196] X. Yuan and R. A. Mayanovic, “An Empirical Study on Raman Peak Fitting and Its Application to Raman Quantitative Research,” *Appl. Spectrosc.*, vol. 71, no. 10, pp. 2325–2338, Oct. 2017, doi: 10.1177/0003702817721527.

[197] R. Gautam, S. Vanga, F. Ariese, and S. Umaphathy, “Review of multidimensional data processing approaches for Raman and infrared spectroscopy,” *EPJ Tech. Instrum.*, vol. 2, no. 1, p. 8, Dec. 2015, doi: 10.1140/epjti/s40485-015-0018-6.

[198] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data*

Mining, Inference, and Prediction. Springer, 2009.

[199] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[200] D. G. Leslie *et al.*, “Identification of Pediatric Brain Neoplasms Using Raman Spectroscopy,” *Pediatr. Neurosurg.*, vol. 48, no. 2, pp. 109–117, 2012, doi: 10.1159/000343285.

[201] X. Chen, M. Wang, and H. Zhang, “The use of classification trees for bioinformatics,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 1, no. 1, pp. 55–63, 2011, doi: 10.1002/widm.14.

[202] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, p. 436, May 2015.

[203] T. Bocklitz, M. Putsche, and A. Niendorf, “A comprehensive study of classification methods for medical diagnosis,” *J Raman Spectrosc*, vol. 40, no. 12, p. 8, 2009.

[204] X. Fan, W. Ming, H. Zeng, Z. Zhang, and H. Lu, “Deep learning-based component identification for the Raman spectra of mixtures,” *Analyst*, vol. 144, no. 5, pp. 1789–1798, 2019, doi: 10.1039/C8AN02212G.

[205] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon, and S. J. Gibson, “Deep convolutional neural networks for Raman spectrum recognition: a unified solution,” *Analyst*, vol. 142, no. 21, pp. 4067–4074, 2017, doi: 10.1039/C7AN01371J.

[206] S. Malek, F. Melgani, and Y. Bazi, “One-dimensional convolutional neural networks for spectroscopic signal regression,” *J. Chemom.*, vol. 32, no. 5, p. e2977, May 2018, doi: 10.1002/cem.2977.

[207] S. L. Neal, “Multivariate Analysis of Mixed Lipid Aggregate Phase Transitions Monitored Using Raman Spectroscopy,” *Appl. Spectrosc.*, vol. 72, no. 1, pp. 102–113, Sep. 2017, doi: 10.1177/0003702817729347.

[208] C. Cui and T. Fearn, “Modern practical convolutional neural networks for multivariate regression: Applications to NIR calibration,” *Chemom. Intell. Lab. Syst.*, vol. 182, pp. 9–20, Nov. 2018, doi: 10.1016/j.chemolab.2018.07.008.

[209] J. Acquarelli, T. van Laarhoven, J. Gerretzen, T. N. Tran, L. M. C. Buydens, and E. Marchiori, “Convolutional neural networks for vibrational spectroscopic data analysis,” *Anal. Chim. Acta*, vol. 954, pp. 22–31, Feb. 2017, doi: 10.1016/j.aca.2016.12.010.

[210] S. Liu and W. Deng, “Very deep convolutional neural network based image classification using small training sample size,” in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, 2015, pp. 730–734, doi: 10.1109/ACPR.2015.7486599.

[211] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[212] C. Szegedy *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[213] D. T. DePaoli, P. Tossou, M. Parent, D. Sauvageau, and D. C. Côté, “Convolutional Neural Networks for Spectroscopic Analysis in Retinal Oximetry,” *Sci. Rep.*, vol. 9, no. 1, p. 11387, Aug.

2019, doi: 10.1038/s41598-019-47621-7.

[214] J. Gu *et al.*, “Recent advances in convolutional neural networks,” *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.

[215] C. Krafft, L. Neudert, T. Simat, and R. Salzer, “Near infrared Raman spectra of human brain lipids,” *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.*, vol. 61, no. 7, pp. 1529–1535, May 2005, doi: 10.1016/j.saa.2004.11.017.

[216] M. Kirsch, G. Schackert, R. Salzer, and C. Krafft, “Raman spectroscopic imaging for in vivo detection of cerebral brain metastases,” *Anal. Bioanal. Chem.*, vol. 398, no. 4, pp. 1707–1713, Oct. 2010, doi: 10.1007/s00216-010-4116-7.

[217] J. M. P. Nascimento and J. M. B. Dias, “Vertex component analysis: a fast algorithm to unmix hyperspectral data,” *IEEE Trans. Geosci. REMOTE Sens.*, vol. 43, no. 4, p. 13, 2005.

[218] M. E. Winter, “N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data,” in *Imaging Spectrometry V*, Oct. 1999, vol. 3753, pp. 266–275, doi: 10.1117/12.366289.

[219] R. E. Kast *et al.*, “Raman molecular imaging of brain frozen tissue sections,” *J. Neurooncol.*, vol. 120, no. 1, pp. 55–62, 2014, doi: 10.1007/s11060-014-1536-9.

[220] T. C. Hollon *et al.*, “Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks,” *Nat. Med.*, vol. 26, no. 1, pp. 52–58, Jan. 2020, doi: 10.1038/s41591-019-0715-9.

[221] P. Larkin, “Chapter 7 - General Outline and Strategies for IR and Raman Spectral Interpretation,” in *Infrared and Raman Spectroscopy*, P. Larkin, Ed. Oxford: Elsevier, 2011, pp. 117–133.

[222] Z. Movasaghi, S. Rehman, and I. U. Rehman, “Raman Spectroscopy of Biological Tissues,” *Appl. Spectrosc. Rev.*, vol. 42, no. 5, pp. 493–541, Sep. 2007, doi: 10.1080/05704920701551530.

[223] M. Köhler, S. Machill, R. Salzer, and C. Krafft, “Characterization of lipid extracts from brain tissue and tumors using Raman spectroscopy and mass spectrometry,” *Anal. Bioanal. Chem.*, vol. 393, no. 5, pp. 1513–1520, Mar. 2009, doi: 10.1007/s00216-008-2592-9.

[224] C. Krafft, S. B. Sobottka, G. Schackert, and R. Salzer, “Near infrared Raman spectroscopic mapping of native brain tissue and intracranial tumors,” *The Analyst*, vol. 130, no. 7, p. 1070, 2005, doi: 10.1039/b419232j.

[225] J. Zhang *et al.*, “Accuracy of Raman spectroscopy in differentiating brain tumor from normal brain tissue,” *Oncotarget*, vol. 8, no. 22, pp. 36824–36831, May 2017, doi: 10.18632/oncotarget.15975.

[226] A. Mizuno, H. Kitajima, K. Kawauchi, S. Muraishi, and Y. Ozaki, “Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours,” *J. Raman Spectrosc.*, vol. 25, no. 1, pp. 25–9, Jan. 1994, doi: 10.1002/jrs.1250250105.

[227] A. Mizuno, T. Hayashi, K. Tashibu, S. Muraishi, K. Kawauchi, and Y. Ozaki, “Near-infrared FT-Raman spectra of the rat brain tissues,” *Neurosci. Lett.*, vol. 141, no. 1, pp. 47–52, Jul. 1992, doi: 10.1016/0304-3940(92)90331-Z.

- [228] C. W. Ong, Z. X. Shen, Y. He, T. Lee, and S. H. Tang, “Raman microspectroscopy of the brain tissues in the substantia nigra and MPTP-induced Parkinson’s disease,” *J. Raman Spectrosc.*, vol. 30, no. 2, pp. 91–96, 1999, doi: 10.1002/(SICI)1097-4555(199902)30:2<91::AID-JRS351>3.0.CO;2-H.
- [229] M. F. Khan, K. Mewes, R. E. Gross, and O. Škrinjar, “Assessment of brain shift related to deep brain stimulation surgery,” *Stereotact. Funct. Neurosurg.*, vol. 86, no. 1, pp. 44–53, 2008.
- [230] U. S. Food and D. Administration, “FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems,” 2018. <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye> (accessed Jun. 31, 2019).
- [231] A. N. Kuzmin, A. Pliss, and P. N. Prasad, “Ramanomics: New omics disciplines using micro raman spectrometry with biomolecular component analysis for molecular profiling of biological structures,” *Biosensors*, vol. 7, no. 4, 2017, doi: 10.3390/bios7040052.
- [232] H. Yan *et al.*, “IDH1 and IDH2 Mutations in Gliomas,” *N. Engl. J. Med.*, vol. 360, no. 8, pp. 765–773, Feb. 2009, doi: 10.1056/NEJMoa0808710.
- [233] C. A. Figueiredo and J. T. Rutka, “Diagnostic imaging: Intraoperative virtual histology,” *Nat. Biomed. Eng.*, vol. 1, no. 2, pp. 1–2, 2017, doi: 10.1038/s41551-017-0033.
- [234] R. Michael, A. Lenferink, G. F. J. M. Vrensen, E. Gelpi, R. I. Barraquer, and C. Otto, “Hyperspectral Raman imaging of neuritic plaques and neurofibrillary tangles in brain tissue from Alzheimer’s disease patients,” *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, 2017, doi: 10.1038/s41598-017-16002-3.
- [235] G. Devitt, K. Howard, A. Mudher, and S. Mahajan, “Raman spectroscopy: An emerging tool in neurodegenerative disease research and diagnosis,” *ACS Chem. Neurosci.*, vol. 9, no. 3, pp. 404–420, 2018, doi: 10.1021/acschemneuro.7b00413.
- [236] L. Zecca, R. Fariello, P. Riederer, D. Sulzer, A. Gatti, and D. Tampellini, “The absolute concentration of nigral neuromelanin, assayed by a new sensitive method, increases throughout the life and is dramatically decreased in Parkinson’s disease,” *FEBS Lett.*, vol. 510, no. 3, pp. 216–220, 2002, doi: 10.1016/S0014-5793(01)03269-0.
- [237] M. G. Bridelli, D. Tampellini, and L. Zecca, “The structure of neuromelanin and its iron binding site studied by infrared spectroscopy,” *FEBS Lett.*, vol. 457, no. 1, pp. 18–22, 1999, doi: 10.1016/S0014-5793(99)01001-7.
- [238] K. Franke, G. Ziegler, S. Klöppel, and C. Gaser, “Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters,” *NeuroImage*, vol. 50, no. 3, pp. 883–892, 2010, doi: 10.1016/j.neuroimage.2010.01.005.
- [239] I. Beheshti, N. Maikusa, and H. Matsuda, “The association between ‘Brain-Age Score’ (BAS) and traditional neuropsychological screening tools in Alzheimer’s disease,” *Brain Behav.*, vol. 8, no. 8, pp. 1–14, 2018, doi: 10.1002/brb3.1020.
- [240] S. Filo *et al.*, “Disentangling molecular alterations from water-content changes in the aging human brain using quantitative MRI,” *Nat. Commun.*, vol. 10, no. 1, p. 3403, 2019, doi: 10.1038/s41467-019-11319-1.

- [241] R. Galli *et al.*, “Intrinsic indicator of photodamage during label-free multiphoton microscopy of cells and tissues,” *PloS One*, vol. 9, no. 10, p. e110295, 2014, doi: 10.1371/journal.pone.0110295.
- [242] M. Krzywinski and N. Altman, “Multiple linear regression,” *Nat. Methods*, vol. 12, no. 12, p. 1103, Dec. 2015.
- [243] V. Cherkassky and Yunqian Ma, “Comparison of loss functions for linear regression,” in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, Jul. 2004, vol. 1, pp. 395–400, doi: 10.1109/IJCNN.2004.1379938.
- [244] P. J. Bickel *et al.*, “Regularization in statistics,” *Test*, vol. 15, no. 2, pp. 271–344, Sep. 2006, doi: 10.1007/BF02607055.
- [245] L. Breiman, “Heuristics of instability and stabilization in model selection,” *Ann Stat.*, vol. 24, no. 6, pp. 2350–2383, Dec. 1996, doi: 10.1214/aos/1032181158.
- [246] A. E. Hoerl and R. W. Kennard, “Ridge Regression: Biased Estimation for Nonorthogonal Problems,” *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970, doi: 10.2307/1267351.
- [247] E.-J. Wagenmakers, M. Lee, T. Lodewyckx, and G. J. Iverson, “Bayesian Versus Frequentist Inference,” in *Bayesian Evaluation of Informative Hypotheses*, H. Hoijtink, I. Klugkist, and P. A. Boelen, Eds. New York, NY: Springer New York, 2008, pp. 181–207.
- [248] R. L. Wasserstein, A. L. Schirm, and N. A. Lazar, “Moving to a World Beyond ‘ $p < 0.05$,’” *Am. Stat.*, vol. 73, no. sup1, pp. 1–19, Mar. 2019, doi: 10.1080/00031305.2019.1583913.
- [249] V. Amrhein, S. Greenland, and B. McShane, “Scientists rise up against statistical significance,” *Nature*, vol. 567, no. 7748, pp. 305–307, Mar. 2019, doi: 10.1038/d41586-019-00857-9.
- [250] L. G. Halsey, “The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum?,” *Biol. Lett.*, vol. 15, no. 5, p. 20190174, May 2019, doi: 10.1098/rsbl.2019.0174.
- [251] J. P. A. Ioannidis, “Why Most Published Research Findings Are False,” *PLoS Med.*, vol. 2, no. 8, p. 6, 2005.
- [252] G. Gigerenzer, “Mindless statistics,” *J. Behav. Exp. Econ. Former. J. Socio-Econ.*, vol. 33, no. 5, pp. 587–606, 2004.
- [253] R. Rosenthal and J. Gaito, “The Interpretation of Levels of Significance by Psychological Researchers,” *J. Psychol.*, vol. 55, no. 1, pp. 33–38, Jan. 1963, doi: 10.1080/00223980.1963.9916596.
- [254] S. Nieuwenhuis, B. U. Forstmann, and E.-J. Wagenmakers, “Erroneous analyses of interactions in neuroscience: a problem of significance,” *Nat. Neurosci.*, vol. 14, no. 9, pp. 1105–1107, Sep. 2011, doi: 10.1038/nn.2886.
- [255] Colquhoun David, “An investigation of the false discovery rate and the misinterpretation of p-values,” *R. Soc. Open Sci.*, vol. 1, no. 3, p. 140216, 2014, doi: 10.1098/rsos.140216.
- [256] A. Gelman *et al.*, *Bayesian Data Analysis*. Chapman and Hall/CRC, 2013.
- [257] J. L. Puga, M. Krzywinski, and N. Altman, “Points of significance: Bayesian statistics,” *Nat. Methods*, vol. 12, pp. 377–378, Apr. 2015, doi: 10.1038/nmeth.3368.

- [258] S. R. Eddy, “What is Bayesian statistics?,” *Nat. Biotechnol.*, vol. 22, no. 9, pp. 1177–1178, Sep. 2004, doi: 10.1038/nbt0904-1177.
- [259] J. K. Kruschke, “Bayesian estimation supersedes the t test,” *J. Exp. Psychol. Gen.*, vol. 142, no. 2, pp. 573–603, 2013, doi: 10.1037/a0029146.
- [260] R. L. Winkler, “The Assessment of Prior Distributions in Bayesian Analysis,” *J. Am. Stat. Assoc.*, vol. 62, no. 319, pp. 776–800, Sep. 1967, doi: 10.1080/01621459.1967.10500894.
- [261] N. P. Lemoine, “Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses,” *Oikos*, vol. 128, no. 7, pp. 912–928, 2019, doi: 10.1111/oik.05985.
- [262] N. G. Polson and V. Sokolov, “Bayesian regularization: From Tikhonov to horseshoe,” *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 11, no. 4, p. e1463, Jul. 2019, doi: 10.1002/wics.1463.
- [263] A. Gelman, D. Simpson, and M. Betancourt, “The Prior Can Often Only Be Understood in the Context of the Likelihood,” *Entropy*, vol. 19, no. 10, p. 555, Oct. 2017, doi: 10.3390/e19100555.
- [264] T. Park and G. Casella, “The Bayesian Lasso,” *J. Am. Stat. Assoc.*, vol. 103, pp. 681–686, Jun. 2008, doi: 10.1198/016214508000000337.
- [265] H. Mallick and N. Yi, “A new Bayesian lasso,” *Stat. Interface*, vol. 7, pp. 571–582, 2014, doi: 10.4310/sii.2014.v7.n4.a12.
- [266] J. Griffin and P. Brown, “Hierarchical Shrinkage Priors for Regression Models,” *Bayesian Anal.*, vol. 12, no. 1, pp. 135–159, Mar. 2017, doi: 10.1214/15-BA990.
- [267] H. Ji and X. S. Liu, “Analyzing ’omics data using hierarchical models,” *Nat. Biotechnol.*, vol. 28, no. 4, pp. 337–340, Apr. 2010, doi: 10.1038/nbt.1619.
- [268] A. Gelman and J. Hill, Eds., “Multilevel regression,” in *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge: Cambridge University Press, 2006, pp. 235–236.
- [269] L. Li and W. Yao, “Fully Bayesian logistic regression with hyper-LASSO priors for high-dimensional feature selection,” *J. Stat. Comput. Simul.*, vol. 88, no. 14, pp. 2827–2851, Sep. 2018, doi: 10.1080/00949655.2018.1490418.
- [270] S. van Erp, D. L. Oberski, and J. Mulder, “Shrinkage priors for Bayesian penalized regression,” 2018, doi: 10.31219/osf.io/cg8fq.
- [271] C. D. Fryar, D. Kruszan-Moran, Q. Gu, and C. L. Ogden, “Mean body weight, weight, waist circumference, and body mass index among adults: United States, 1999–2000 through 2015–2016,” *Natl. Health Stat. Rep.*, no. 122, Dec. 2018, [Online]. Available: <https://stacks.cdc.gov/view/cdc/61430>.
- [272] D. van Ravenzwaaij, P. Cassey, and S. D. Brown, “A simple introduction to Markov Chain Monte–Carlo sampling,” *Psychon. Bull. Rev.*, vol. 25, no. 1, pp. 143–154, Feb. 2018, doi: 10.3758/s13423-016-1015-8.
- [273] S. Chib and E. Greenberg, “Understanding the Metropolis-Hastings Algorithm,” *Am. Stat.*, vol. 49, no. 4, pp. 327–335, Nov. 1995, doi: 10.1080/00031305.1995.10476177.
- [274] S. Sharma, “Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy,” *Annu. Rev. Astron. Astrophys.*, vol. 55, no. 1, pp. 213–259, Aug. 2017, doi:

10.1146/annurev-astro-082214-122339.

[275] B. Carpenter *et al.*, “Stan: A Probabilistic Programming Language,” *J. Stat. Softw.*, vol. 76, no. 1, pp. 1–32, Jan. 2017, doi: 10.18637/jss.v076.i01.

[276] A. S. Haka *et al.*, “In vivo Margin Assessment during Partial Mastectomy Breast Surgery Using Raman Spectroscopy,” *Cancer Res*, vol. 66, no. 6, pp. 3317–3322, 2006.

[277] M. S. Bergholt, K. Lin, W. Zheng, D. P. C. Lau, and Z. Huang, “In vivo, real-time, transnasal, image- guided Raman endoscopy: defining spectral properties in the nasopharynx and larynx,” *J. Biomed. Opt.*, vol. 17, p. 8, 2012.

[278] H. Lui, J. Zhao, D. McLean, and H. Zeng, “Real-time Raman Spectroscopy for In Vivo Skin Cancer Diagnosis,” *Cancer Res.*, p. 11.

[279] J. L. Griffin and J. P. Shockcor, “Metabolic profiles of cancer cells.,” *Nat. Rev. Cancer*, vol. 4, no. 7, pp. 551–61, 2004.

[280] G. Reifenberger, H.-G. Wirsching, C. B. Knobbe-Thomsen, and M. Weller, “Advances in the molecular genetics of gliomas — implications for classification and therapy,” *Nat. Rev. Clin. Oncol.*, vol. 14, no. 7, pp. 434–452, Dec. 2016, doi: 10.1038/nrclinonc.2016.204.

[281] T. Wyss-Coray, “Ageing, neurodegeneration and brain rejuvenation,” *Nature*, vol. 539, no. 7628, pp. 180–186, Nov. 2016, doi: 10.1038/nature20411.

[282] R. D. Riley *et al.*, “Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes,” *Stat. Med.*, vol. 38, no. 7, pp. 1276–1296, Mar. 2019, doi: 10.1002/sim.7992.

[283] T. van der Ploeg, P. C. Austin, and E. W. Steyerberg, “Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints,” *BMC Med. Res. Methodol.*, vol. 14, no. 1, p. 137, Dec. 2014, doi: 10.1186/1471-2288-14-137.

[284] R. K. Reddy and R. Bhargava, “Chemometric Methods for Biomedical Raman Spectroscopy and Imaging,” in *Emerging Raman Applications and Techniques in Biomedical and Pharmaceutical Fields*, P. Matousek and M. D. Morris, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 179–213.

[285] M. Busch *et al.*, “Influence of brain tumors on the MR spectra of healthy brain tissue,” *Magn. Reson. Med.*, vol. 65, no. 1, pp. 18–27, 2011, doi: 10.1002/mrm.22612.

[286] S. J. Choquette, E. S. Etz, W. S. Hurst, D. H. Blackburn, and S. D. Leigh, “Relative Intensity Correction of Raman Spectrometers: NIST SRMs 2241 through 2243 for 785 nm, 532 nm, and 488 nm/514.5 nm Excitation,” *Appl. Spectrosc.*, vol. 61, no. 2, pp. 117–129, Feb. 2007, doi: 10.1366/000370207779947585.

[287] R. Perez-Pueyo, M. J. Soneira, and S. Ruiz-Moreno, “Morphology-Based Automated Baseline Removal for Raman Spectra of Artistic Pigments,” *Appl. Spectrosc.*, vol. 64, no. 6, pp. 595–600, Jun. 2010, doi: 10.1366/000370210791414281.

[288] H. nath Banerjee and L. Zhang, “Deciphering the finger Prints of Brain Cancer Astrocytoma in comparison to Astrocytes by using near infrared Raman Spectroscopy,” *Mol. Cell. Biochem.*, vol. 295, no. 1–2, pp. 237–240, Jan. 2007, doi: 10.1007/s11010-006-9278-4.

[289] S. van Erp, D. L. Oberski, and J. Mulder, “Shrinkage priors for Bayesian penalized

- regression,” *J. Math. Psychol.*, vol. 89, pp. 31–50, Apr. 2019, doi: 10.1016/j.jmp.2018.12.004.
- [290] S. Kim, S. H. Lee, S. Y. Min, K. M. Byun, and S. Y. Lee, “Dual-modal cancer detection based on optical pH sensing and Raman spectroscopy,” *J. Biomed. Opt.*, vol. 22, no. 10, pp. 1–6, 2017, doi: 10.1117/1.JBO.22.10.105002.
- [291] J. Lever, M. Krzywinski, and N. Altman, “Principal component analysis,” *Nat. Methods*, vol. 14, no. 7, pp. 641–642, Jul. 2017, doi: 10.1038/nmeth.4346.
- [292] J. D. Gelder, K. D. Gussem, P. Vandenabeele, and L. Moens, “Reference database of Raman spectra of biological molecules,” *J Raman Spectrosc*, vol. 38, no. 9, p. 15, 2007, doi: 10.1002/jrs.1734.
- [293] B. Ogretmen, “Sphingolipid metabolism in cancer signalling and therapy,” *Nat. Rev. Cancer*, vol. 18, no. 1, pp. 33–50, Nov. 2017, doi: 10.1038/nrc.2017.96.
- [294] A. Rygula, K. Majzner, K. M. Marzec, A. Kaczor, M. Pilarczyk, and M. Baranska, “Raman spectroscopy of proteins: a review,” *J. Raman Spectrosc.*, vol. 44, no. 8, pp. 1061–1076, Aug. 2013, doi: 10.1002/jrs.4335.
- [295] I. Guyon and A. Elisseeff, “An Introduction to Variable and Feature Selection,” *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.
- [296] Y. Zhang, J. N. Hofmann, M. P. Purdue, S. Lin, and S. Biswas, “Logistic Bayesian LASSO for genetic association analysis of data from complex sampling designs,” *J. Hum. Genet.*, vol. 62, pp. 819–829, Apr. 2017, doi: 10.1038/jhg.2017.43.
- [297] A. Huang, S. Xu, and X. Cai, “Empirical Bayesian LASSO-logistic regression for multiple binary trait locus mapping,” *BMC Genet.*, vol. 14, no. 1, p. 5, 2013, doi: 10.1186/1471-2156-14-5.
- [298] L. Pasanen, L. Holmström, and M. J. Sillanpää, “Bayesian LASSO, Scale Space and Decision Making in Association Genetics,” *PLOS ONE*, vol. 10, no. 4, p. e0120017, Apr. 2015, doi: 10.1371/journal.pone.0120017.
- [299] J. Surmacki, B. Brozek-Pluska, R. Kordek, and H. Abramczyk, “The lipid-reactive oxygen species phenotype of breast cancer. Raman spectroscopy and mapping, PCA and PLSDA for invasive ductal carcinoma and invasive lobular carcinoma. Molecular tumorigenic mechanisms beyond Warburg effect,” *The Analyst*, vol. 140, no. 7, pp. 2121–2133, Apr. 2015, doi: 10.1039/c4an01876a.
- [300] S. Robel, B. Berninger, and M. Götz, “The stem cell potential of glia: lessons from reactive gliosis,” *Nat. Rev. Neurosci.*, vol. 12, no. 2, pp. 88–104, Feb. 2011, doi: 10.1038/nrn2978.
- [301] P. H. Huang, A. M. Xu, and F. M. White, “Oncogenic EGFR Signaling Networks in Glioma,” *Sci. Signal.*, vol. 2, no. 87, p. 14, Sep. 2009.
- [302] P. Lu, V. M. Weaver, and Z. Werb, “The extracellular matrix: A dynamic niche in cancer progression,” *J Cell Biol*, vol. 196, no. 4, pp. 395–406, Feb. 2012, doi: 10.1083/jcb.201102147.
- [303] A. Monteiro, R. Hill, G. Pilkington, and P. Madureira, “The Role of Hypoxia in Glioblastoma Invasion,” *Cells*, vol. 6, no. 4, p. 45, Nov. 2017, doi: 10.3390/cells6040045.
- [304] M. G. Martinez, A. J. Bullock, S. MacNeil, and I. U. Rehman, “Characterisation of structural changes in collagen with Raman spectroscopy,” *Appl. Spectrosc. Rev.*, pp. 1–34, Jan.

2019, doi: 10.1080/05704928.2018.1506799.

[305] T. T. Nguyen, C. Gobinet, J. Feru, S. B. -Pasco, M. Manfait, and O. Piot, “Characterization of Type I and IV Collagens by Raman Microspectroscopy: Identification of Spectral Markers of the Dermo-Epidermal Junction,” *Spectrosc. Int. J.*, vol. 27, pp. 421–427, 2012, doi: 10.1155/2012/686183.

[306] H. Chow and K. Herrup, “Genomic integrity and the ageing brain,” *Nat. Rev. Neurosci.*, vol. 16, no. 11, pp. 672–684, Nov. 2015, doi: 10.1038/nrn4020.

[307] C. Choi *et al.*, “2-hydroxyglutarate detection by magnetic resonance spectroscopy in IDH-mutated patients with gliomas,” *Nat. Med.*, vol. 18, no. 4, pp. 624–629, Apr. 2012, doi: 10.1038/nm.2682.

[308] M. N. Kinalwa, E. W. Blanch, and A. J. Doig, “Accurate Determination of Protein Secondary Structure Content from Raman and Raman Optical Activity Spectra,” *Anal. Chem.*, vol. 82, no. 15, pp. 6347–6349, Aug. 2010, doi: 10.1021/ac101334h.

[309] I. Anna, P. Bartosz, P. Lech, and A. Halina, “Novel strategies of Raman imaging for brain tumor research,” *Oncotarget*, vol. 8, no. 49, pp. 85290–85310, Oct. 2017, doi: 10.18632/oncotarget.19668.

[310] Y. Zhou, C. Liu, K. Zhu, B. Wu, X. Yu, and R. R. Alfano, “4 - Visible Resonance Raman Spectroscopy in Human Brain Tissues,” in *Neurophotonics and Biomedical Spectroscopy*, R. R. Alfano and L. Shi, Eds. Elsevier, 2019, pp. 65–106.

[311] L. S. Payne and P. H. Huang, “The Pathobiology of Collagens in Glioma,” *Mol. Cancer Res.*, vol. 11, no. 10, pp. 1129–1140, Oct. 2013, doi: 10.1158/1541-7786.MCR-13-0236.

[312] H. Lemos, L. Huang, G. C. Prendergast, and A. L. Mellor, “Immune control by amino acid catabolism during tumorigenesis and therapy,” *Nat. Rev. Cancer*, vol. 19, no. 3, pp. 162–175, Mar. 2019, doi: 10.1038/s41568-019-0106-z.

[313] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

[314] J. S. Hawe, F. J. Theis, and M. Heinig, “Inferring Interaction Networks From Multi-Omics Data,” *Front. Genet.*, vol. 10, p. 535, 2019, doi: 10.3389/fgene.2019.00535.

[315] P. Grabowski and J. Rappsilber, “A Primer on Data Analytics in Functional Genomics: How to Move from Data to Insight?,” *Trends Biochem. Sci.*, vol. 44, no. 1, pp. 21–32, Jan. 2019, doi: 10.1016/j.tibs.2018.10.010.

[316] B. A. Richards *et al.*, “A deep learning framework for neuroscience,” *Nat. Neurosci.*, vol. 22, no. 11, pp. 1761–1770, Nov. 2019, doi: 10.1038/s41593-019-0520-2.

[317] K. R. Merikangas and N. Risch, “Genomic Priorities and Public Health,” *Science*, vol. 302, no. 5645, pp. 599–601, Oct. 2003, doi: 10.1126/science.1091468.

[318] W. H. Rogowski, S. D. Grosse, and M. J. Khoury, “Challenges of translating genetic tests into clinical and public health practice,” *Nat. Rev. Genet.*, vol. 10, no. 7, pp. 489–495, Jul. 2009, doi: 10.1038/nrg2606.

[319] J. I. Hare, T. Lammers, M. B. Ashford, S. Puri, G. Storm, and S. T. Barry, “Challenges and strategies in anti-cancer nanomedicine development: An industry perspective,” *Adv. Drug Deliv.*

Rev., vol. 108, pp. 25–38, Jan. 2017, doi: 10.1016/j.addr.2016.04.025.

[320] C. Zhang and Y. Ma, *Ensemble machine learning: methods and applications*. Springer, 2012.