scientific reports

OPEN

Preliminary study demonstrating cancer cells detection at the margins of whole glioblastoma specimens with Raman spectroscopy imaging

François Daoust[1,2], Frédérick Dallaire[1,2], Hugo Tavera[1,2], Katherine Ember[1,2], Marie-Christine Guiot[3], Kevin Petrecca[4] & Frederic Leblond[1,2,5]✉

Intraoperative Raman spectroscopy uses near-infrared laser light to gain molecular information without causing damage. It can be used in vivo or ex vivo without exogenous contrast agents. Clinically, the technique was primarily used with machine learning for in situ tumor detection with fiberoptics probes analyzing tissue at sub-millimeter scales one point at the time. Here we report the development of a whole-specimen spectroscopic imaging system designed to detect cancer cells at the margins of surgical specimens. The system has a field of view covering a square area of side one centimeter with a pixel size of a quarter of a millimeter . First, a tumor detection model was developed from data acquired using a point-probe in 24 glioblastoma patients that had a detection sensitivity of 90% and a specificity of 95%. That model was then used to produce cancer prediction maps of nine glioblastoma specimens from five patients with validation based on histopathology analyses. The results preliminarily demonstrate the instrument was able to detect tissue areas associated with cancer cells from the Raman peaks associated with the amino acids phenylalanine and tryptophan as well as the relative concentration of lipids and proteins linked with deformations of the CH_2 and CH_3 bonds.

It is estimated that approximately 45 million surgical procedures are performed annually, with surgical intervention representing a primary treatment option for numerous cancers[1]. The objective of surgical oncology is to excise tumors and guarantee a secure margin of surrounding tissue while preserving as much healthy tissue and organ function as is feasible. The removal of the primary tumor remains a fundamental aspect of surgical oncology and cancer treatment[2]. The original cancerous lesion is excised, and in the case of many cancers, the surgical margins are examined by histopathology after surgery to ascertain the absence of residual cancer cells. A residual tumor is indicative of a positive surgical margin and often necessitates a second surgical procedure to reduce the risk of recurrence[3,4]. However, the analysis of tissue samples by histopathology is a challenging process that requires significant time and expertise[5,6]. The involvement of trained pathologists, who are in high demand, further complicates the intraoperative assessment of surgical margins[7]. Currently, there is a lack of decision support tools available during surgical procedures[8], particularly those that can provide real-time feedback without the need for the administration of exogenous contrast agents or the time-consuming process of tissue staining[9].

In the context of surgical interventions for brain cancers, it is of particular importance to achieve a balance between complete tumor removal and the minimization of impact on normal tissue, with a view to improving survival and quality of life[10–12]. This is particularly crucial in the context of gliomas, which are characterized by the presence of infiltrative cancer cells, i.e. cells that have extended beyond the boundaries of the primary tumor mass into the surrounding brain tissue. This infiltrative behavior is a defining characteristic of gliomas of all World Health Organization (WHO) grades, including Grade II-III oligodendrogliomas and astrocytomas, as well as glioblastomas (WHO Grade IV). As a result, complete surgical removal is a challenging endeavor. Gliomas constitute approximately 24.5 percent of all primary brain tumors[13] and 80% of all primary malignant

[1]Polytechnique Montréal, Montreal, Canada. [2]Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montreal, Canada. [3]Division of Neuropathology, Department of Pathology, Montreal Neurological Institute-Hospital, McGill University, Montreal, Canada. [4]Montreal Neurological Institute-Hospital, McGill University, Montreal, Canada. [5]Institut du Cancer de Montréal, Montreal, Canada. ✉email: frederic.leblond@polymtl.ca

brain tumors[14]. Glioblastomas are the most aggressive form, representing approximately 49% of all malignant brain tumors[13]. The time from surgery to tumor recurrence for glioblastomas (progression-free survival) and patient life expectancy are inversely correlated with the volume of tumor remaining after surgery[15]. Moreover, the presence of residual tumor may necessitate further surgical intervention[16,17]. Nevertheless, resections that extend beyond the margins of the tumor into the adjacent brain can result in neurological deficits, which may have an adverse impact on the quality of life and overall survival of the patient[18]. In the case of brain cancers, post-surgery margin assessment is done through magnetic resonance imaging (MRI) scans. In instances where residual tumor is identified, a follow-up surgery may be necessary. This approach is costly, time-consuming and carries an additional risk to the patient[19,20]. New surgical tools are needed to help detect tumor tissue during a procedure.

Near-infrared Raman spectroscopy is a potential solution allowing real-time tissue analyses and tumor margin detection[21,22]. It is an analytical technique that uses laser light to gain molecular information about a sample. The application of Raman spectroscopy in an intraoperative setting does not require the administration of an exogenous contrast agent and, when used in pathology for surgical specimen inspection, does not require tissue staining. The inelastic (Raman) scattering of laser light by a sample is dependent on vibrational states and thus the structure and bonding of molecules. As molecular composition of biological samples can alter in disease states such as cancer and ischemia, samples may be classified based on their Raman spectral profiles. The energetic differences between the incident near-infrared light and the inelastically scattered light can be plotted on the x-axis to form a spectrum with the intensity of Raman scattering at a particular wavenumber given on the y-axis. The peak position of the x-axis gives information about molecular identity whilst the peak height gives information about relative molecular concentrations. A single Raman peak may have contributions from different types of biomolecules if they have molecular bonds with similar vibrational states.

A Raman spectroscopy system has already been developed for in situ detection of brain tumors during surgery[23,24]. This system has proven able to detect glioblastoma, meningioma, and brain metastases intraoperatively with greater than 90% accuracy based on data from 67 neurosurgical cases[25]. Whilst this was achieved rapidly and in a reagent-free way, this single-point device gains molecular information from a relatively small portion of tissue (0.25 cm^2) with no imaging capability. To gain biomolecular information from larger tissue areas and to create a molecular map of tumor margins, a line-scanning imaging system was developed. This allowed illumination and molecular analysis of an area of 1 cm^2[26,27]. Here, we present a pilot study with that system demonstrating detection of residual cancer cells on the surface of surgical specimens. At the core of the device is a whole-specimen Raman spectroscopy unit for real-time reagent-free molecular analysis. This relies on a tumor detection machine learning model which was trained in vivo on the pre-existing single-point system data, from 24 patients undergoing glioblastoma surgery[25]. Equipped with this model, the large field-of-view Raman imaging device was tested on brain specimens from five glioblastoma patients, including tumor and non-tumor tissue. The results from that technique were validated based on histopathology analyses and could potentially be generalized to other cancers to allow intra-operative margins inspection in surgical specimens[21,28–30].

Results

Hyperspectral Raman imaging instrument

A Raman spectroscopy imaging system was developed to detect residual cancer cells in the surgical margin (Fig. 1). The device was tested on surgically excised tumor specimens, including specimens from the bulk of the tumor and samples from its edge (peripheral samples). The instrument included a portable imaging probe with a field-of-view (FOV) of 1 cm^2[26,27]. The cart-mounted system was portable and allowed its deployment directly into the operating room for testing in fresh brain specimens from glioblastoma patients. Tumor was resected during neurosurgery and specimens were immediately analyzed by Raman imaging. Spectral images and co-located brightfield images were acquired for each specimen, both with a square-shaped field-of-view of 1 cm by 1 cm. The Raman images contained 1,440 pixels (40 by 36 pixels, each of size 200 µm by 250 µm). Each pixel was associated with one individual Raman spectrum of spectral resolution 8 cm^{-1} with a spectral domain ranging from 400 to 1800 cm^{-1}.

Each spectrum contained biomolecular tissue information from more than 30 distinguishable peaks of known biomolecular origin. All peaks were associated with different molecular bonds. Intensity of Raman scattering was proportional to the concentration of molecules within a sample[31]. Therefore, the intensity of one biomolecular peak relative to all other peaks in a single spectrum was used as a semi-quantitative readout of the concentration of different families of biomolecules. The molecular origin of the peaks could be traced back to protein backbones, aromatic amino acids (phenylalanine, tryptophane, tyrosine), different classes of lipids, and the structural units of nucleic acids[32,33].

Tumor detection machine learning model

A glioblastoma detection machine learning model (herein referred to as the *Tumor detection model*) was created using data from a well-validated “single-point system”. The single-point system illuminated a spot of tissue of 0.5 mm diameter circular area and obtained a single Raman spectrum from the region under interrogation (Fig. 1). Spectroscopic data was acquired independently using this system in the scope of a multi-center clinical study[25]. Briefly, the tumor detection model was developed using support vector machines (SVM) and five-fold cross-validation (Methods). It was trained/validated and tested using 352 spectra from 24 patients (glioblastoma: $n = 248$, normal brain: $n = 104$). The spectral data associated with normal brain was on average composed of an equal proportion of measurements sampling normal grey matter, normal white matter, or a mixture of both. All spectra were acquired in vivo during tumor resection and prior to tissue removal.

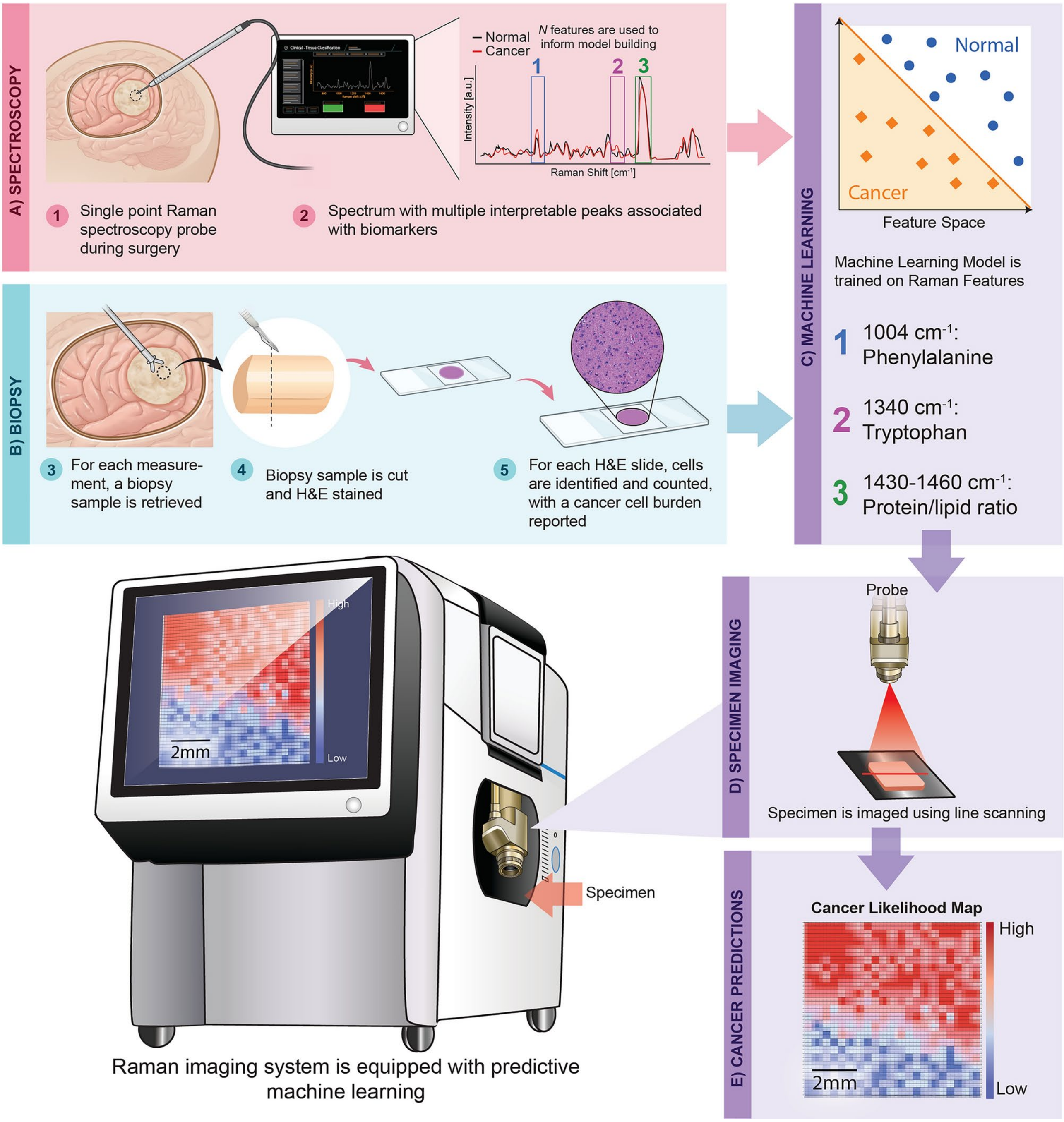


Fig. 1. Graphical depiction of the workflow leading to cancer prediction maps of specimens resected from glioblastoma patients that were imaged using the Raman spectroscopy imaging system. (**A**) The experimental workflow showcases how glioblastoma and normal brain Raman spectra were collected. Spectral fingerprint measurements were acquired using a single-point probe during neurosurgery. (**B**) Biopsy workflow showing how histopathology data was linked with each spectral measurement. A biopsy was recovered immediately following the Raman measurement, the specimen was prepared and then annotated by a neuropathologist. (**C**) The combination of Raman spectra and their histological labels allowed the creation of machine learning classifiers based on a select number of differentiating Raman features. The phenylalanine band (1004 cm^{-1}), the tryptophan band (1340 cm^{-1}) and the protein/lipid band (1430–1460 cm^{-1}) were used to create a cancer (i.e., glioblastoma) detection model. In (**D**), schematic depiction of the imaging system using line scanning to image freshly resected brain specimen. (**E**) The cancer prediction model was applied to each pixel on the Raman images to recover a cancer likelihood map that can detect tumor tissue (red) and areas of non-tumoral brain (blue).

The resulting predictive model was integrated into the data processing and operating software of the large field-of-view Raman imaging instrument, hereafter named the "Raman imaging system". Each time a Raman image was acquired from an excised brain specimen using the Raman imaging system, the tumor detection model was directly applied to each individual spectrum within that image. Each image pixel had an output consisting of a predictive metric between 0 and 1. Values closer to 0 were more likely taken from tissue associated with *Normal brain*, while values closer to 1 had spectral features consistent with *Tumor* tissue. Each image was hence a map of the *Cancer Likelihood* metric, showing the spatial distribution within specimens in terms of *Normal brain* or *Tumor* likelihood ("Methods").

Spectral features originating from three Raman bands were selected through the machine learning model development process: 1004 cm^{-1}, 1340 cm^{-1}, and 1430–1460 cm^{-1} (Fig. 1). These peaks were chosen because they are part of the five key peaks identified by Ember et al. and they did not exhibit deviations likely due to intrinsic differences between Raman systems. Other studies have demonstrated the phenotypic relevance of these spectral regions and explained their biological origin. The Raman peak at 1004 cm^{-1} is related to the concentration of the amino acid phenylalanine[32,34,35] and the inelastic scattering signal in the spectral region around 1340 cm^{-1} is associated with the protein collagen and the amino acid tryptophan[32,36]. The peak covering the 1430–1460 cm^{-1} spectral domain contains information on the relative concentration of lipids and proteins and is related to the deformations of the CH_2 and CH_3 molecular bonds[32,35]. The glioblastoma model previously published by our group used features from the same three bands but also included features from the band around 1300 cm^{-1} associated with CH_2 saturated bonds of lipids. However, that band was excluded from the machine learning model development process (along with the bands around 1620 cm^{-1} and 1660 cm^{-1}) because of observed inter-systems discrepancies (see explanations below) in those regions. However, this minimally impacted the cross-validation performance of the tumor detection SVM classifier. The performance of the model used as part of the work here was 91% accuracy, 90% sensitivity and 95% specificity, while the performance reported in our other work[25] was 91% accuracy, 91% sensitivity and 91% specificity.

Patient specimens and histology

Between one and three tumor specimens were analyzed from each of five glioblastoma patients undergoing tumor resection surgery at the Montreal Neurological Institute and Hospital. The samples were prepared and imaged using the Raman imaging system ("Methods"). In total, nine specimens were analyzed, each labeled with a code consisting of a patient identification number (P5 to P9) and a set of letters (SPA, SPB or SPC) identifying each individual specimen acquired from the same patient (Table 1). A brightfield image of each specimen was acquired (Fig. 2: first column) and a digitized H&E histology map generated for each specimen (Fig. 2: third column), which was co-located with the Raman data by a trained neuropathologist ("Methods").

The *Cancer Likelihood* images obtained from the Raman imaging system were produced (Fig. 2: second column) as well as a *Cellular Density Map* from each H&E image. The latter quantified the number of cells for each imaged pixel (Fig. 2, fourth column). Cellularity colormaps were produced by counting the number of nuclei per unit area in the stained histology section ("Methods"). Each specimen was classified by the neuropathologist as belonging to one of four histological categories. Specimens P7 SPA-SPC were labeled *All Tumor* because they were composed of only tumor tissue. Specimens P5 SPA-B and P6 SPA were labeled *All Tumor with Extensive Necrosis* because necrosis was observed over approximately 20% of the surface area. The remaining three specimens were either associated with a low density of cancer cells spread throughout the tissue (*Low Cellularity Tumor*: P9 SPA-B) or consisted of normal brain mixed with tissue having a low density of cancer cells (*Normal brain with Infiltration*: P8 SPB).

Biomolecular information content and inter-system spectral discrepancies

The Raman imaging dataset consisted of 11,361 spectra acquired from 9 specimens and the tumor detection model was applied to each spectrum. All spectra that had a *Cancer Likelihood* value higher than 0.8 were assigned to the *Tumor* category, whilst all other spectra were assigned to the *Normal brain* category. Overall, the classifier

Specimen ID	Sex	Age	Histological label	Cancer likelihood metric (avg ± SD)	Cellular density (cells/m^2) (avg ± SD)
P5 SPA	F	83	All tumor with extensive necrosis	0.88 ± 0.19	600 ± 600
P5 SPB	F	83	All tumor with extensive necrosis	0.96 ± 0.10	1000 ± 600
P6 SPA	M	83	All tumor with extensive necrosis	0.98 ± 0.07	2500 ± 1200
P7 SPA	F	67	All tumor	0.98 ± 0.08	3000 ± 800
P7 SPB	F	67	All tumor	0.99 ± 0.04	3400 ± 700
P7 SPC	F	67	All tumor	0.98 ± 0.09	2400 ± 800
P8 SPB	M	61	Normal brain with infiltration	0.30 ± 0.19	900 ± 300
P9 SPA	F	59	Low cellularity tumor	0.52 ± 0.30	1500 ± 300
P9 SPB	F	59	Low cellularity tumor	0.56 ± 0.28	1400 ± 300

Table 1. Information associated with the patients (sex, age) and pathology of the glioblastoma specimens. P5 to P9 refer to patient identification numbers, while SP*X* (*X* : A, B or C) refers to different specimens from the same patient. Other information include the histopathological assessment for each specimen, the average and standard deviation (SD) of the *Cancer likelihood* metric computed for all pixels within the Raman images, and the corresponding values of the cellular density in units of cells per mm^2.

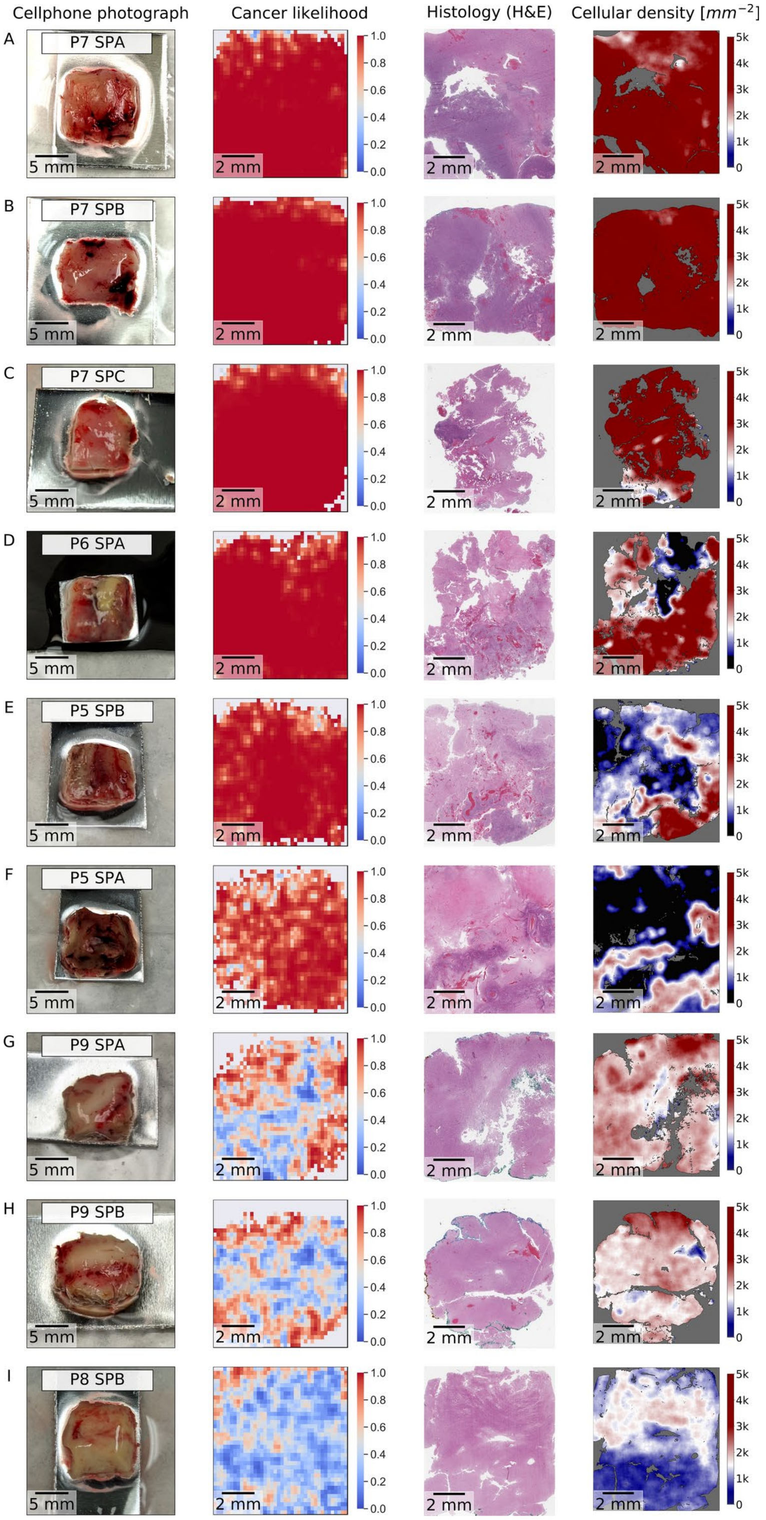


Fig. 2. Glioblastoma specimens with corresponding photograph (first column), machine learning predictions from a tumor model trained using in vivo point probe Raman spectroscopy data (second column), haematoxylin and eosin (H&E) stained digital scan (third column) and cellular density map (fourth column). Cancer likelihood predictions gave a score from 0 (blue) to 1 (red) where 1 corresponded to the highest probability of cancer cells while 0 corresponded to the lowest probability of cancer (normal brain prediction). Cellular density maps were produced by counting the number of nuclei per surface area (per mm^2) with black areas labelled as necrotic tissue.

predicted that 8,071 pixels (71% of all image data points) were associated with *Tumor*, and 3,290 pixels (29% of all image data points) were associated with *Normal brain*.

A visual comparison of the spectra acquired with the Raman imaging and single-point systems demonstrated that the detected biomolecular features were mostly consistent between the two systems (Fig. 3). The mean *Tumor* spectrum from the imaging system showed similarities with the mean glioblastoma spectrum from the single-point system dataset (Fig. 3a). The mean *Normal brain* imaging spectrum also showed similarities with the mean normal brain spectrum taken using the single-point system (Fig. 3b). The spectra acquired with both systems prior to background removal were also shown (Supplementary Fig. 2). For both systems, the prominent Raman peaks observed for the glioblastoma and *Normal brain* categories were at 1004, 1064, 1090, 1128, 1300, 1340, 1430–1460, 1620 and 1660 cm^{-1} (Fig. 3: red, green, and blue highlights), each with its own molecular interpretation (Supplemental Table 1). The consistency of spectra acquired with the Raman imaging system and the single-point probe was studied and quantified as part of another study using synthetic material and ex vivo porcine brain tissue[37].

The Raman peaks around 1064 cm^{-1} and 1090 cm^{-1} were associated with the carbon skeletal stretching of lipids[38] while the inelastic scattering signal at 1128 cm^{-1} was associated with the CH_2 wagging mode found in lipids[38]. The peak at 1004 cm^{-1} was linked with the phenylalanine tissue content and the Raman band around 1300 cm^{-1} was linked to CH_2 twist and wag as well as with the Amide III band of proteins[38]. Similarly, the band at 1660 cm^{-1} was associated with the Amide I band of proteins, more specifically sourced from the carbon–

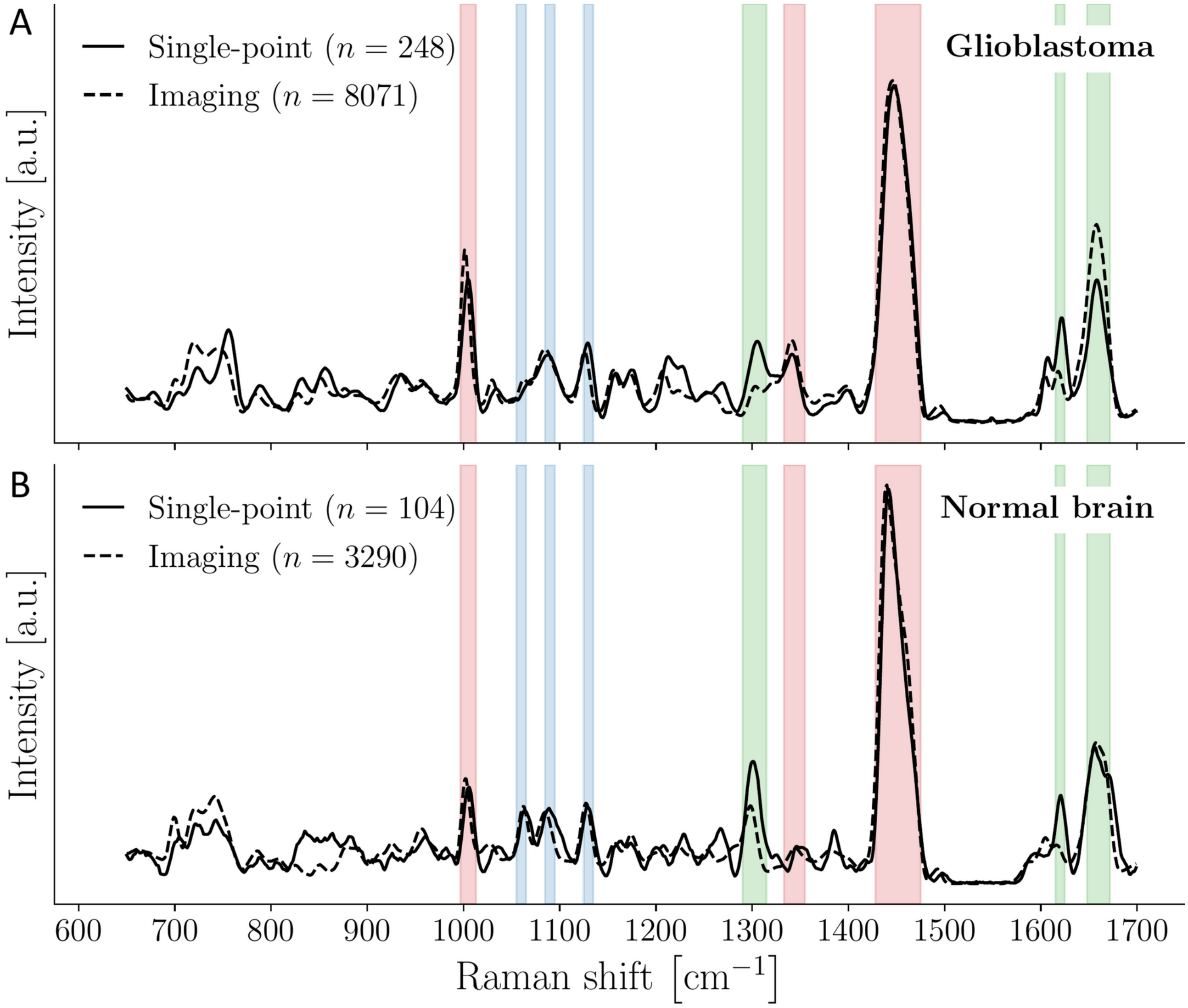


Fig. 3. Average Raman spectra from the single-point probe (solid lines) and the imaging system (dashed lines) datasets: (**A**) glioblastoma, (**B**) normal brain. Bands highlighted in red represent spectral regions containing the features used by the machine learning tumor model. Bands highlighted in green showcase regions where systematic differences between the point-probe and the imaging system were observed. Bands highlighted in blue are regions showing similar signatures between the two systems but that were not exploited by the tumor detection model.

oxygen double bond in peptide bonds[38]. The band around 1340 cm^{-1} was linked with tryptophan and CH_2/CH_3 deformations in lipids and protein. The band around 1620 cm^{-1} was associated with contributions from the aromatic amino acids: tyrosine, tryptophane and phenylalanine[39–41]. The band located in the 1430–1460 cm^{-1} range included several vibrational modes involving carbon and hydrogen linked to protein and lipid content[41,42].

Whilst the spectra detected with both systems showed the same Raman bands, differences in peak intensities were observed between the two imaging configurations. These mostly occurred within the protein bands around 1300, 1620 and 1660 cm^{-1} (Fig. 3: red highlights). The single-point measurements were taken in vivo, whilst the imaging measurements were taken ex vivo. Therefore, the differences could be attributed to tissue deterioration (e.g., denaturation of the proteins) from environmental changes. These differences may also be linked to the automated application of the fluorescence background removal algorithm ("Methods"), which may have caused overfitting under some peaks. Importantly, none of the spectral features utilized by the tumor detection machine learning model were associated with the bands that had been identified as due to intrinsic differences between the Raman systems (Fig. 3: green highlights). Statistics associated with the intensity of individual Raman bands provides a more quantitative assessment of the overlap between spectra acquired with the two systems (Fig. 4).

Consistency between machine learning predictions and histology

The Raman imaging system was able to correctly detect the different types of tumorous and non-tumorous tissue that were sampled by the neurosurgeon. The predictive power of the tumor detection machine learning model was evaluated using the co-located histology and cellular density maps that were here considered as the as gold standard for histopathological assessment (Fig. 2). Specimens P7 SPA, P7 SPB and P7 SPC were all classified as pure tumor based on histopathology analyses. The cellular density maps exhibited hypercellularity, i.e. cell density higher than in normal brain. Pathology analyses revealed multiple other features consistent with glioblastoma, including microvascular proliferation and pseudopalisading necrosis. This latter feature is

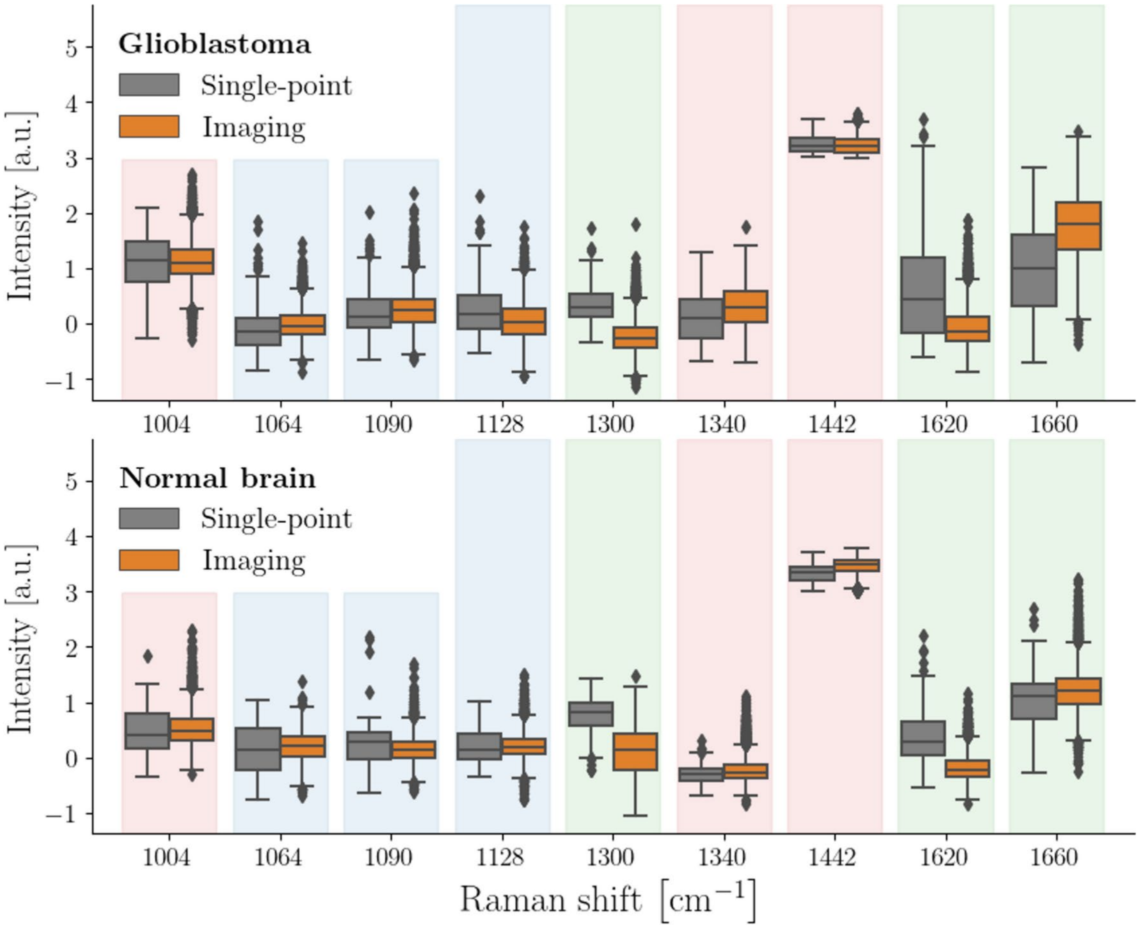


Fig. 4. Differences between the intensity of inelastic scattering bands for the spectra shown in Fig. 3 that were acquired with the single-point probe system and the Raman imaging system. Boxplots are shown that were associated with the Raman band intensity of all spectra at 1004, 1064, 1090, 1128, 1300, 1340, 1442, 1620 and 1660 cm^{-1}. The boxplots displayed the dataset minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum values. The box spans Q1 to Q3, showing the interquartile range (IQR), while the line inside marks the median (Q2). The whiskers extend to the minimum and maximum within 1.5 times the IQR, and any points beyond are plotted individually as outliers.

a hallmark of glioblastoma associated with cells migrating from the necrotic tumor core. The Raman imaging system correctly predicted that these specimens contained tumor tissue across the whole field-of-view. The average *Cancer Likelihood* was high at 0.98 with a standard deviation (0.07) that was consistent with a mostly homogenous specimen.

Specimens P5 SPA, P5 SPB and P6 SPA were tumor tissue that contained necrotic tissue. They all exhibited hypercellularity within some regions, whilst other regions had additional histological features typical of glioblastoma, including microvascular proliferation and extensive areas of necrosis. Necrotic areas were associated with a low cell density, confirmed by the absence of cellular architecture and nuclei on the histology maps. These histological features suggest that the specimens were taken in the vicinity of the necrotic core of the tumor. Despite the difference in morphology and cell density compared to the P7 tumor features, the Raman imaging system correctly predicted that most of the area of these specimens were tumor tissue. However, the specimens had a slightly lower average *Cancer Likelihood* metric and were more heterogenous when compared to P7 samples. The average *Cancer Likelihood* for these samples was 0.94 with a standard deviation of 0.13. This slightly lower *Cancer Likelihood* could be because the surface area of the specimens was associated with a larger fraction of necrotic tissue. Of note, the tumor detection model was trained from data acquired with the single-point system which did not contain measurements associated with necrotic tissue. This may explain why this model did not appear to be able to distinguish tissue containing necrotic regions from tissue containing cancer cells (Fig. 2).

The histological features associated with specimens P9 SPA and P9 SPB were different compared with specimens from P5, P6 and P7. They exhibited hypercellularity in some areas but only limited microvascular proliferation with no areas associated with necrosis nor pseudopalisading necrosis. In addition, a reduction in cellularity compared to other tumor specimens was observed in both P9 SPA and P9 SPB that was associated with an organized tissue architecture characteristic of normal brain. These features were consistent with specimens resected at the periphery of the tumor, i.e. towards the outer edge the MRI-enhancing ring of the glioblastoma. Specifically, the Raman images were consistent with what would be expected from samples with a mixture of normal and cancer cells. The average *Cancer Likelihood* for these two specimens was 0.54 with a standard deviation of 0.29, consistent with a highly heterogenous specimen. The Raman images showed areas of predicted tumor mixed with normal brain areas.

The last specimen considered –P8 SPB– mostly consisted of normal brain, with histology features consistent with an organized cellular architecture. It also had a cellularity that was low compared to all other specimens. The specimen only had a few signs of tissue infiltration with cancer cells, low mitotic activity, and a low level of nuclear atypia. It had none of the distinguishing features found in glioblastoma such as microvascular proliferation and necrosis. The average *Cancer Likelihood* was low at 0.30 with a standard deviation consistent with moderate tissue heterogeneity 0.19.

Biomolecular alterations between different anatomical regions

The Raman imaging system used different biomolecular features to identify tumor and non-tumor pathologies. An analysis was performed to determine precisely which biomolecular features were characteristic of each type of pathology. Firstly, specimens were divided into four categories based on their suspected anatomical origin and the histology analyses. The 3,464 spectra from the samples P5 SPA, P5 SPB, P6 SPA were labeled *Tumor Core*. Specimens P7 SPA, P7 SPB and P7 SPC were likely associated with the MRI-enhancing ring of the tumor with the 4,079 spectra labeled *Tumor Interphase*. Spectra from P8 SPB, P9 SPA and P9 SPB that had a *Cancer Likelihood* metric larger than 0.8 ($n=528$) were labeled *Tumor periphery – Cancer* while all other spectra were labeled *Tumor periphery – Normal brain* ($n=3{,}290$). Of note, those spectra were not all histologically confirmed to be normal brain, but they were all predicted to belong to the normal brain category based on the tumor detection machine learning model. Despite this limitation, these spectra were used as baseline to assess biomolecular changes between specimens having different histology features.

The average spectrum for each tissue category was computed along with the inter-measurement standard deviation within each class (Fig. 5a). The difference between the spectra for *Tumor core*, *Tumor interphase* and *Tumor periphery – Cancer* compared to the baseline normal brain spectrum (*Tumor periphery – Normal brain*) was also computed (Fig. 5b). These *difference spectra* allowed identification of the main Raman-predicted biomolecular changes between different tumor tissues and tissue that had a more organized cellular architecture. A quantitative comparison of normal brain spectra was made with samples labeled *Tumor Core* (i.e. *All Tumor with Extensive Necrosis*) and *Tumor interphase* (i.e. *All Tumor*) that were acquired with the Raman imaging system (Supplemental Fig. 1). This was done to assess spectral differences potentially orginating from necrotic tissue.

There were no significant differences observed within the Raman bands used by the tumor detection machine learning model (1004 cm^{-1}, 1340 cm^{-1}, 1442 cm^{-1}) between specimens from the *Tumor core*—which had extensive areas of necrosis—and those from the *Tumor interphase*. This meant the model was not able to distinguish cancer from necrotic tissue. However, there were spectral regions that qualitatively showed spectral differences between *Tumor core* and *Tumor interphase*. Those regions included, for example, the low wavenumber range below 800 cm^{-1} and molecular bonds associated with phospholipids and nucleic acids in the peaks around 1100 cm^{-1}.

There were however significant differences between *Tumor core/Tumor interphase*, *Tumor periphery – Cancer* and *Tumor periphery – Normal brain* for the Raman bands used by the tumor detection model. All observed changes were consistent with a gradual decrease of the ratio of lipids to proteins, as well as the proteins and amino acid concentration, when moving away from the *Tumor core/Tumor interphase* region towards areas showing fewer distinguishing features of cancer. This is manifested by an increase – in tumor tissue – of the intensity of the Raman peaks associated with phenylalanine (1004 cm^{-1}) and tryptophan (1340 cm^{-1}), combined

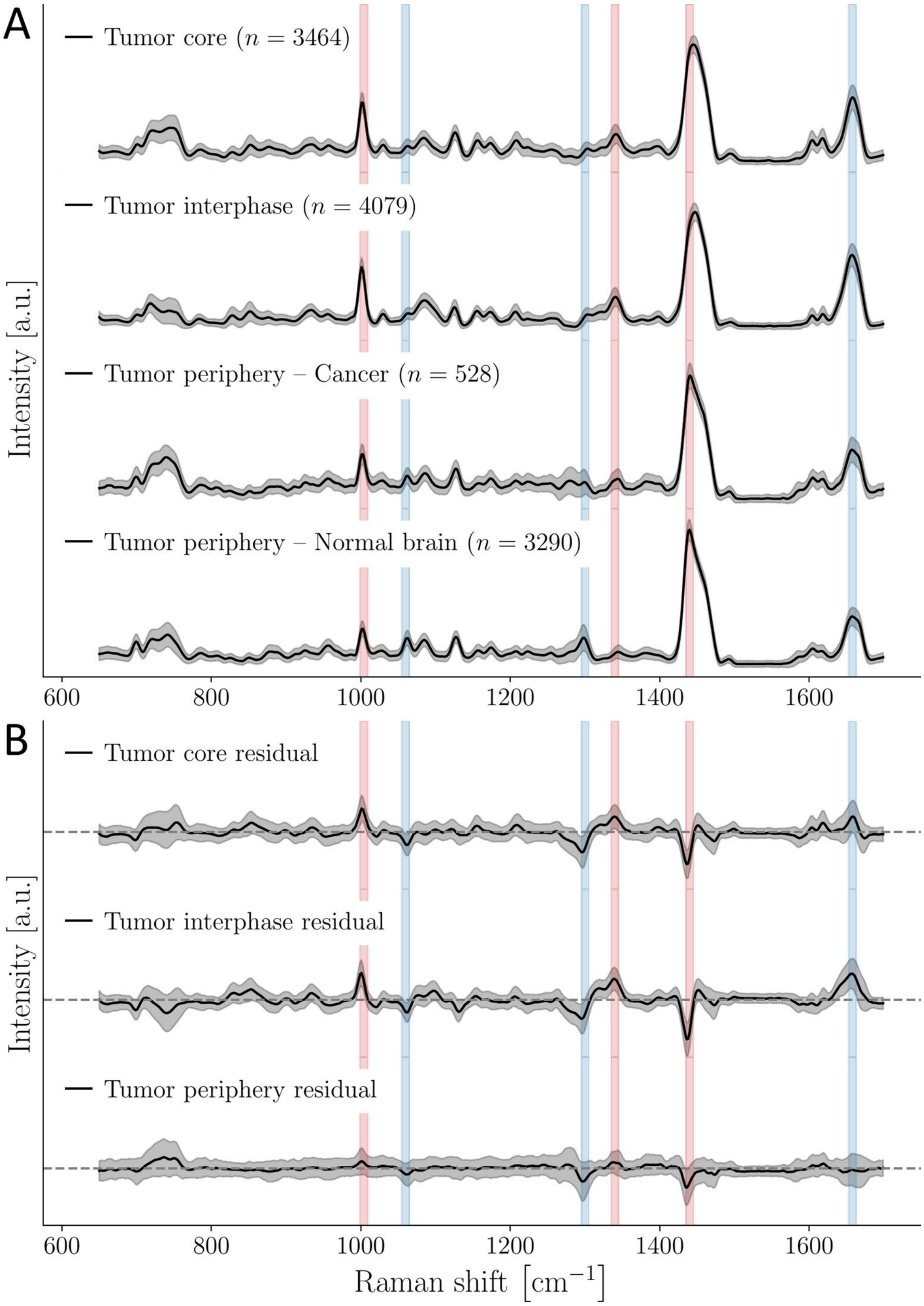


Fig. 5. (**A**) Average Raman spectra (solid line) and inter-measurements standard deviation (grey shading) of specimens that were predicted as *Tumor* (first three graphs) or *Normal brain* (fourth graph). The first graph shows spectra associated with specimens with a mixture of cancer cells and necrotic tissue (labeled *Tumor core*). The second graph shows spectra associated with the tumor interphase that were mostly cancer with no necrotic tissue. The third and fourth graphs show spectra from mixed normal-cancer tissues areas that were labeled *Tumor periphery*. The third graph shows spectra that were predicted to belong to the *Tumor* category while the fourth graph shows spectra predicted as *Normal brain*. (**B**) Difference between the Raman spectra predicted as *Tumor* from specimens labeled *Tumor core*, *Tumor interphase*, or *Tumor periphery* and the spectra predicted as *Normal brain*. The standard deviation is shown in grey shading. The dashed grey line corresponds to 0. The spectral features used in the tumor prediction model are highlighted in red. Bands highlighted in blue show other features where differences are observed between tumor and normal brain but that were not exploited by the machine learning model.

with a decrease of lipid to protein ratio computed from the 1430–1460 cm^{-1} band. Other protein bands showed a marked decrease between tumor and normal brain, including at 1300 cm^{-1} and 1660 cm^{-1}.

Discussion

Predicting the presence of cancer cells from large field-of-view Raman imaging was achieved, allowing sub-millimeter resolution maps of glioblastoma specimens to be produced. The agreement between histology and the Raman-predicted *Cancer Likelihood* maps suggested the method could effectively assess margins in a real-world intraoperative context. The device detected tumor tissue from three molecular bands. Relative to normal brain, glioblastoma exhibited a lower lipid to protein ratio based on spectral features extracted from the spectral range 1430–1460 cm^{-1}. Glioblastoma also showed an overexpression of phenylalanine (1004 cm^{-1}) as well as collagen and tryptophan (1340 cm^{-1}). These trends were consistent across specimens, including pure tumor associated with a high density of cancer cells, specimens presenting with a mixture of high-density cancer cell clusters and necrosis, as well as non-necrotic tissue associated with a low density of cancer cells.

A limitation of this study was the accuracy of the spatial registration between histopathology images (H&E maps, cellular density images) and the images acquired with the Raman imaging system, i.e. the *Cancer Likelihood* maps. The study employed a spatial registration method using inked specimen borders to allow an approximate geometrical matching between the Raman images and the histology maps. However, the specimen preparation protocol could have altered its shape, limiting the accuracy of the affine registration with the corresponding Raman images. The method allowed regions of high *Cancer Likelihood* to be consistently associated with tumor tissue and regions of low *Cancer Likelihood* with normal brain. However, precise one-to-one pixel matching with histology images was not feasible because of the spatial registration inaccuracies. Consequently, it was not possible to compute accuracy, sensitivity, and specificity performance metrics for the Raman imaging system on a pixel basis. To improve pixel matching in future studies, physical registration markers (e.g., ink spots, needles) should be placed on the specimen after Raman imaging. This approach could improve the accuracy of affine image registration with histology, allowing to compute tumor model sensitivity and specificity when applied to Raman imaging data.

The tumor detection machine learning model used to produce the *Cancer Likelihood* maps was trained using a different instrument, in the scope of an independent multicentric study. This model was not trained to distinguish areas associated with a high density of cancer cells from regions with necrotic areas. However, this did not imply there were no spectral differences between those different tissue types. Rather, this meant the subset of spectral features used to train the tumor detection model were present in tumor tissue regardless of the presence of necrosis. In fact, spectral changes were visually detected between tumor specimens that included necrotic tissue and those that did not (Supplemental Fig. 1). Tissues containing necrotic areas showed elevated inelastic scattering signal at 1128, 1442 and 1620 cm^{-1}, associated with lipids and protein and a decrease of the peaks at 1090, 1340 and 1660 cm^{-1} which are predominantly associated with tryptophan and amino acids. This suggested that a machine learning model could eventually be trained specifically to detect necrosis.

The latter observation could have profound implications for the future of technology. Other clinical studies could be designed aimed at developing more complex machine learning models detecting more subtle phenotypes. For example, detecting the invasive gradient of glioblastoma is important in glioma surgery[43]. A model could be developed aiming at detecting subtle tissue changes associated with cellular proliferation beyond the MRI-enhancing ring of a glioblastoma. This was not attempted as part of the current study because detecting invasive glioblastoma cells with high accuracy is difficult with conventional histology techniques. There is, in fact, no gold standard for this because there are no definitive histological or protein markers of wildtype glioblastoma cells. Using standard technologies, it is thus not possible to count invasive glioblastoma cells with clinical grade accuracy, unlike astrocytoma and oligodendrogliomas in which single cells can be detected with an antibody to the IDHR132H mutation[44].

Unlocking new applications will require improvements in system design leading to enhanced signal-to-noise ratio, i.e. *super pixels* associated with less stochastic noise. Here the system was only able to acquire high signal-to-noise ratio for a subset of Raman bands, principally the three bands used to train the tumor detection model. This was mainly due to the 30-min limit to imaging time for integration into the workflow. Thirty minutes of imaging time per specimen aligns with the duration of frozen section pathology, which is commonly used during glioma surgery and typically takes 15 to 30 min per specimen[45]. There were also light losses due to a 2 m coherent imaging bundle used in the detection path of the system[26,27]. The imaging bundle was imposed by prior design constraints requiring the system to acquire signal at-a-distance through a hand-held imaging probe in situ within the surgical cavity. This aspect of the design could be relaxed for ex vivo margins assessment application, improving light collection efficiency by at least a factor 10 × . Other aspects reducing imaging times – that would not require any modifications of the hardware – could include the reduction of the spatial resolution along the scanning axis of the system by a factor 2, with limited impact in the ability of the system to detect the whole surface area of the specimen. A re-design of the system could thus allow the imaging time achieved in the current work to be reduced to 90 s. Further, keeping the same imaging time as in this current work – 30 min on average – would lead to a signal-to-noise ratio increase of 4.5X, improving overall signal quality to increase the number of spectral features available to develop new more complex machine learning models.

One of the most enticing aspects of the current work is the application of the system to other types of cancer surgeries in organs where promising Raman spectroscopy studies were already published. For example, a workflow could be developed in breast surgical oncology allowing lumpectomy specimens to be imaged during a surgical excision procedure. If residual tumor was detected on the specimen, the surgeon could then resect extra tissue. This could improve outcomes by reducing disease recurrence whilst minimizing damage to healthy tissue. The same approach could apply in radical prostatectomy procedures either performed laparoscopically or using a robotic platform[46]. Pilot clinical studies have already been performed using a hand-held single-point Raman spectroscopy probe. These demonstrate tumor detection in multiple organs including breast[47], prostate[48–50], lung[51], the ovaries, and endometrium[52].

Methods

Raman imaging system

Raman images of the specimens were acquired using the system published by Daoust et al.[27]. Briefly, the imaging system was cart-mounted and equipped with an imaging probe connected to two flexible coherent fiber optics imaging bundles (one for excitation and one for detection) of length ~ 2 m. Excitation patterns (linear laser light distributions) generated by the system were conveyed through the excitation bundle and probe to the sample. Spectroscopic and white light images were captured through the detection bundle. The imaging probe, originally designed for in situ tissue characterization within the surgical cavity, was affixed within a light-tight optical enclosure for ex vivo imaging (Fig. 1). The system was operated in non-contact mode with a numerical aperture of 0.22 and a working distance of 40 mm between the front lens and the specimen surface. Images consisted of 40 lines projected from a 785 nm laser (5W, Innovative Photonic Solutions, USA) that was scanned across a square field-of-view of side dimension 1 cm. Each line had a 200 μm thickness along the scanning axis and 10 mm along the transverse direction. Images were formed by scanning the line with 0.25 mm steps. Spatial binning was applied along the second dimension on the charge-coupled device (CCD) sensor (1,024 × 256 pixels array with 26 μm pixels) resulting in super pixels that covered a square area with a side dimension of 0.25 mm. The resulting dimension of each Raman image was 40 *super pixels* along the scanning axis (i.e., 40 lines), 36 *super pixels* along the transverse axis and 1,024 light intensity values along the wavenumber (cm^{-1}) axis. Throughout the manuscript, "pixels" refers to these *super pixels*. Monte Carlo simulations and experimental validations by Akbarzadeh et al. report that for a similar sensing geometry, absorption and reduced scattering coefficients consistent with human brain ($\mu_a = 0.01\ mm^{-1}$ and $\mu_s' = 1\ mm^{-1}$)[53] and 785 nm excitation wavelength, the sensing depth of Raman signal was no more than 600 μm[54].

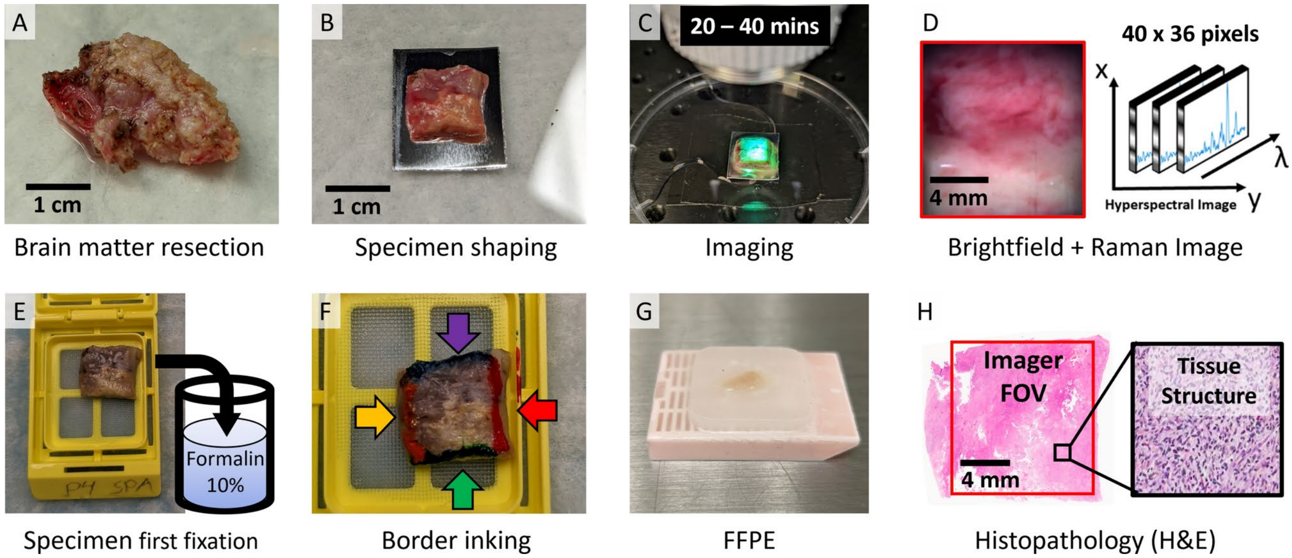
Spectra were produced through spectral separation of light using a high-resolution diffraction grating spectrometer (EH Model, EmVision, USA) coupled to a cooled CCD camera (Newton 920, Andor Technology, Oxford Instruments, USA). The camera sensor was cooled to -60 °C prior to each measurement. The spectral range of each spectrum was 400 to 2100 cm^{-1} (810 to 940 nm) with a spectral resolution of approximately 8 cm^{-1}. Laser light power at the specimen was 905 mW when measured for an entire line, resulting in an average intensity of 22.6 W/cm^2. The exposure time for each line ranged from 4 to 15 s. These exposure levels did not lead to histological features on H&E-stained sections consistent with tissue damage. The later was optimized to maximize usage of the CCD sensor dynamical range while avoiding saturation. This was done to maximize the signal-to-noise ratio (SNR). The upper bound of 15 s per line was set to keep the total scan time for 40 lines within 10 min. Each measurement consisted of a signal averaged over $N = 4$ to 10 repetitions to further improve SNR. Tests in brain tissue showed that beyond $N = 10$ repetitions, SNR gains were minimal. The total imaging time per specimen averaged 30 min, with a range from 20 to 40 min, aligning with the time typically required for frozen section pathology (approximately 30 min). These acquisition parameters resulted in an integration time of 30 to 60 s per *super pixel*. The system allowed brightfield images to be acquired that were collocated with the hyperspectral Raman images.

Patient selection, specimen preparation and histology analyses

Five patients were recruited that had a diagnosis of glioblastoma and were scheduled for a tumor resection procedure at the Montreal Neurological Institute and Hospital (McGill University Health Center). All participants signed a written Information and Consent Form. Sufficient information was provided to allow each patient to make an informed decision about their participation in the study.

A procedure was developed that allowed specimens to be harvested and prepared before being imaged with the Raman spectroscopy imaging system (Fig. 6). Specimens were recovered during surgery that were destined for biological waste and that had a volume ranging between 1 and 4 cm^3. Each sample was immediately soaked in phosphate buffered solution (312–651, Thermo Scientific, USA) and placed in a sterile petri dish. The specimens were then sliced using a sterile razor blade to form two or three smaller flat samples of approximate size 10 mm × 10 mm × 3 mm. Each specimen was then placed on an aluminum surface (Miro5011, Anomet, Canada) because this material has limited inelastic scattering signal in the 400 to 2100 cm^{-1} wavenumber range. A phosphate buffer solution was regularly sprayed onto the specimen during the imaging session between image repetitions to keep it moistened, avoid dehydration and regulate temperature. The buffer solution did not induce noticeable changes in the Raman tissue signature.

Raman imaging was performed as described above in a vacant area next to the patient during surgery. After the imaging session was completed, a procedure was applied to each specimen to ensure its spatial orientation was registered. This was necessary to ensure Raman and white light images could be spatially collocated with histopathology images. Each specimen was first immersed in paraformaldehyde 10% for 18 to 24 h. The borders of each square-shaped specimen were then inked following three steps: 1) the fixing agent acetic acid 5% was applied on all borders, 2) four different colors of ink were applied on each side of the specimen (top: blue, right: red, bottom: green, left: yellow or orange), 3) the ink was left to dry for five minutes. Each specimen was then placed in a micro-tissue pathology cassette and submerged for another 24 h in fresh paraformaldehyde 10% before undergoing paraffin-embedding. Slices were then obtained that covered the whole top surface of the resulting formalin-fixed paraffin-embedded (FFPE) specimen. Each slice –with an area consistent with the field-of-view of the Raman imaging system– was stained with hematoxylin and eosin (H&E) and digitally scanned. Recombining all images resulted in a histopathology map co-located with Raman and white light images that was analyzed by a neuropathologist.



Brain matter resection

Specimen shaping

Imaging

Brightfield + Raman Image

Specimen first fixation

Border inking

FFPE

Histopathology (H&E)

Fig. 6. Raman imaging workflow that was developed for brain tissue preparation prior to Raman spectroscopy imaging and histopathology analyses. (**A**) Brain matter was resected during surgery and (**B**) shaped into square specimens of dimension 10 mm by 10 mm by 3 mm. (**C**) Spectroscopic imaging of the specimen took between 20 and 40 min and resulted in a 40 by 36 hyperspectral Raman image. (**D**) A collocated brightfield image was acquired to be used for spatial registration with histology. (**E**) Specimens underwent a first 18 to 24 h fixation and (**F**) their border were inked to ensure spatial orientation information was retained. (**G**) Specimens finally underwent standard tissue preparation consisting of formaldehyde fixation, paraffin-embedding (FFPE). (**H**) The paraffin block was sliced: each slice was stained with hematoxylin and eosin (H&E) and scanned digitally. The digital scans were oriented with the Raman images using the ink visible in the digital scan as a fiducial marker.

Cellular density maps were formed from the digital histology scans of the H&E-stained brain specimens slices using the open-source bioimage analysis software QuPath 0.5.0 [22]. QuPath's*Cell Detection* function was first used to locate all cells within the digital scan using the default values set by the software. Then, the *Create Density Maps* function was used to produce a cellular density image, i.e. a lower resolution image where each pixel value in the image corresponds to the density of cells surrounding the pixel. Each pixel value was determined by counting the number of cells surrounding the pixel within a radius of 282 µm and dividing this cell count by the surface area considered (i.e., 0.25 mm^2). All cellular density values of pixels located in blank areas of the histology scan (i.e. empty areas with no visible cells or tissue) were set to 0. Python 3.11.3 with Scikit-image 0.20.0 were finally used to generate the cellular density maps with colors showcasing areas of necrosis in black, higher cellularity (> 1500 cells/mm^2) in red and lower cellularity (< 1500 cells/mm^2) in blue.

Spectral pre-processing

Data analysis and spectral pre-processing were done using Python 3.9.5 with Scikit-Learn 1.2.2[55] in combination with code from a public repository[56]. The following data pre-processing steps were applied to all spectra associated with the hyperspectral images: (1) filtering using a two-dimensional median filter to remove occasional cosmic ray events, (2) wavenumber correction using a reference acetaminophen measurement, (3) averaging over N repeat measurements, (4) instrument response correction from a NIST 785 nm Raman standard measurement, (5) removing the system intrinsic background signal using measurements with no sample within the imaging focal plane, (6) baseline subtraction using the *BubbleFill* algorithm[56] with a minimum 'bubble' diameter of 40 cm^{-1}, (7) truncation of spectra to the range 650–1700 cm^{-1} and interpolation in 1 cm^{-1} steps, (8) standard normal variate (SNV) normalization. The data from step 2, step 4 and step 5 were associated with images covering the same field-of-view as the specimen images. This allowed calibration to be achieved on a pixel-by-pixel basis. Further, spectral smoothing was done by averaging each pixel with its closest neighbours: 2 pixels for the corners of the image, 3 pixels for the borders, and 4 pixels otherwise.

Finally, a metric was computed for each spectrum that allowed to evaluate their spectral quality[56]. The quality factor (QF) was used to automatically dismiss poor quality spectra that were plagued with high levels of stochastic noise. Only spectra from images with QF ≥ 0.4 were retained in this study. The QF was defined as the average signed squared intensity:

$$QF = \frac{1}{L}\sum_i^L sgn\left(r_i\right) \cdot r_i^2,$$

Where L is the number of spectral bins composing each SNV-normalized spectrum and r_i is the i^{th} spectral element.

Tumor detection machine learning model

The spectral pre-processing procedure reduced the number of spectra retained per Raman image from 1,440 to 1,150 Raman spectra, i.e. a loss of 21% of the whole dataset. A tumor detection machine learning model was produced using data acquired independently with a different system. That model was applied to each spectrum associated with hyperspectral data acquired using the Raman imaging system. Briefly, inelastic scattering spectra of glioblastoma and normal brain were recovered from a previous study that was conducted using a single-point Raman spectroscopy probe instrument[25]. In situ spectroscopic measurements were acquired in vivo with a sterilizable handheld point probe system. One biopsy specimen was extracted along with each spectroscopic measurement ensuring spatial collocation. The probe interrogated a surface area of approximately 0.25 mm^2 and the size of the biopsy specimen was approximately 1–2 mm. The specimens were formalin-fixed paraffin-embedded (FFPE). A section of the resulting FFPE tissue was sliced, stained (H&E) and digitized for histology analyses. Acquisition parameters with the probe system were: (1) excitation wavelength: 785 nm, (2) exposure time: 100 ms, (3) number of repeat measurements: $N = 20$, (4) laser intensity on the brain: 75 mW, (5) spectral range: 400–2100 cm^{-1} (810–940 nm). Each spectrum was labelled either as glioblastoma or *Normal brain* by a neuropathologist. The dataset consisted of 104 normal brain spectra and 248 glioblastoma spectra from a total of 24 patients.

A tumor detection classifier was produced from a dimensionally reduced dataset that was limited to the spectral fingerprint region previously shown to have the most contrast between glioblastoma and the *Normal brain* category[57]: (1) the phenylalanine band at 1004 cm^{-1}, (2) the tryptophan band region located at 1340 cm^{-1}, 3) the lipid-protein band centered around 1430–1460 cm^{-1}. Specifically, all spectral features were associated with the spectral bins in the regions 997–1013 cm^{-1}, 1333–1355 cm^{-1} and 1428–1475 cm^{-1}, resulting in a total of 85 spectral features (i.e. intensity values from SNV-normalized Raman spectra). The feature space was further reduced to less than 20 individual features using a support vector machine (SVM) with a linear kernel and L1 regularization (L1-SVM)[58]. This Lasso regression technique limited the number of features to the ones presenting the most significant contribution to inter-class separation.

The tumor detection model was then trained from the subset of features using a linear SVM algorithm. Hyperparameters –both for the feature selection procedure described above and for the classification algorithms– were optimized using a grid search approach. The L1-SVM regularization parameter C was varied between 0.005 and 0.05, the reduced number of features between 5 and 20, and the classification regularization parameter C between 0.01 and 0.5. A class weight parameter corresponding to the normal/tumor class ratio was used to account for the unbalance between the two classes and avoid classification biases. The performance of the trained model was assessed through five-fold cross-validation where model predictions were compared to the histopathological predictions. A receiver operating characteristic (ROC) analysis was used to determine the set of hyperparameters yielding the best performance. The model trained using this set of hyperparameters corresponded to the tumor detection model applied to each *super pixel* with the hyperspectral Raman images.

Cancer likelihood maps of Raman images

After all spectral pre-processing steps were applied, the spectral features used by the tumor detection model were computed for all super pixels contained within the Raman images. The model was then applied to the pixels associated with all specimen images. The output of the model was then used to generate the *Cancer Likelihood* maps. The output was a probability between 0 and 1 with values close to 0 associated with *Normal brain* and those closer to 1 to *Tumor*. A threshold optimizing sensitivity and specificity was obtained from a ROC analysis and used to classify the pixels in binary categories: spectra with a probability lower than this threshold (0.8) were classified as *Normal brain*, the other ones as *Tumor*. This threshold yielded to the best performance of the tumor detection model during training based on cross-validation (i.e. 91% accuracy, 90% sensitivity and 95% specificity).

The average of the *Cancer Likelihood* metric and the cellular density metric were computed over the surface of each specimen, along with their fluctuations measured by the standard deviation within each image (Table 1). All other information (white-light image, digitized H&E histology map, cellular density image) was displayed, ranking the specimens in decreasing order of the average cellular density computed from the cellular density maps (Fig. 2). This allowed differences in spatial distribution of cancer cells across different specimens to be more easily visualized and compared.

Data availability

The main data supporting the findings of this study are available within the Article and its Supplementary Information. The raw data generated during the study is too large to be publicly shared, yet it is available for research purposes from the corresponding author upon reasonable request.

Code availability

Source code used in this work is available for non-commercial purposes from the corresponding author on request.

Received: 30 March 2024; Accepted: 16 January 2025

Referencess

1. NIHR Global Health Research Unit on Global Surgery. Quality and outcomes in global cancer surgery: Protocol for a multicentre, international, prospective cohort study (GlobalSurg 3). *BMJ Open* **9**, e026646 (2019).

2. Phillips, J. D. & Wong, S. L. Patient-reported outcomes in surgical oncology: An overview of instruments and scores. *Ann. Surg. Oncol.* **27**, 45–53 (2020).
3. Houvenaeghel, G. et al. Positive or close margins: Reoperation rate and second conservative resection or total mastectomy?. *Cancer Manag. Res.* **11**, 2507–2516 (2019).
4. Schmitz, A. K. et al. Second-look surgery after pediatric brain tumor resection: Single center analysis of morbidity and volumetric efficacy. *Brain Spine* **2**, 100865 (2022).
5. Buesa, R. J. Histology: A unique area of the medical laboratory. *Ann. Diagn. Pathol.* **11**, 137–141 (2007).
6. Titford, M. & Bowman, B. What may the future hold for histotechnologists?. *Lab. Med.* **43**, e5–e10 (2012).
7. Metter, D. M., Colgan, T. J., Leung, S. T., Timmons, C. F. & Park, J. Y. Trends in the US and Canadian pathologist workforces from 2007 to 2017. *JAMA Netw. Open* **2**, e194337–e194337 (2019).
8. Dell'Oglio, P. et al. Precision surgery. *Am. J. Nucl. Med. Mol. Imaging* **12**, 74–80 (2022).
9. Ogrinc, N., Saudemont, P., Takats, Z., Salzet, M. & Fournier, I. Cancer surgery 2.0: Guidance by real-time molecular technologies. *Trends Mol. Med.* **27**, 602–615 (2021).
10. Petrecca, K., Guiot, M.-C., Panet-Raymond, V. & Souhami, L. Failure pattern following complete resection plus radiotherapy and temozolomide is at the resection margin in patients with glioblastoma. *J. Neurooncol.* **111**, 19–23 (2013).
11. Rahman, M. et al. The effects of new or worsened postoperative neurological deficits on survival of patients with glioblastoma. *J. Neurosurg. JNS* **127**, 123–131 (2017).
12. Han, S. J. et al. Subcortical stimulation mapping of descending motor pathways for perirolandic gliomas: Assessment of morbidity and functional outcome in 702 cases. *J. Neurosurg.* **131**, 201–208 (2019).
13. Schaff, L. R. & Mellinghoff, I. K. Glioblastoma and other primary brain malignancies in adults: A review. *JAMA* **329**, 574–587 (2023).
14. Pellerino, A., Caccese, M., Padovan, M., Cerretti, G. & Lombardi, G. Epidemiology, risk factors, and prognostic factors of gliomas. *Clin. Transl. Imaging* **10**, 467–475 (2022).
15. Brown, T. J. et al. Association of the extent of resection with survival in glioblastoma: A systematic review and meta-analysis. *JAMA Oncol.* **2**, 1460–1469 (2016).
16. Weller, M., Cloughesy, T., Perry, J. R. & Wick, W. Standards of care for treatment of recurrent glioblastoma-are we there yet?. *Neuro-Oncology* **15**, 4–27. https://doi.org/10.1093/neuonc/nos273 (2013).
17. Tosoni, A., Franceschi, E., Poggi, R. & Brandes, A. A. Relapsed glioblastoma: Treatment strategies for initial and subsequent recurrences. *Curr. Treat. Opt. Oncol.* **17**, 49 (2016).
18. Sanai, N. & Berger, M. S. Glioma extent of resection and its impact on patient outcome. *Neurosurgery* **62**, 753–766 (2008).
19. Young, R. M., Jamshidi, A., Davis, G. & Sherman, J. H. Current trends in the surgical management and treatment of adult glioblastoma. *Ann. Transl. Med.* **3**, (2015).
20. Eidel, O. et al. Tumor infiltration in enhancing and non-enhancing parts of glioblastoma: A correlation with histopathology. *PLoS ONE* **12**, e0169292 (2017).
21. DePaoli, D. et al. Rise of Raman spectroscopy in neurosurgery: A review. *J. Biomed. Opt.* **25**, 050901 (2020).
22. Murugappan, S., Tofail, S. A. M. & Thorat, N. D. Raman spectroscopy: A tool for molecular fingerprinting of brain cancer. *ACS Omega* **8**, 27845–27861 (2023).
23. Desroches, J. et al. Characterization of a Raman spectroscopy probe system for intraoperative brain tissue classification. *Biomed. Opt. Express* **6**, 2380–2397 (2015).
24. Jermyn, M. et al. Intraoperative brain cancer detection with Raman spectroscopy in humans. *Sci. Transl. Med.* **7**, 27419 (2015).
25. Ember, K. et al. In situ brain tumor detection Using the Raman spectroscopy sentry system: Results of a multicenter study. *Sci. Rep.* **14**, 13309 (2024).
26. Daoust, F. et al. Handheld macroscopic Raman spectroscopy imaging instrument for machine-learning-based molecular tissue margins characterization. *J. Biomed. Opt.* **26**, 1–18 (2021).
27. Daoust, F. et al. A clinical Raman spectroscopy imaging system and safety requirements for in situ intraoperative tissue characterization. *Analyst* **148**, 1991–2001 (2023).
28. Zúñiga, W. C. et al. Raman spectroscopy for rapid evaluation of surgical margins during breast cancer lumpectomy. *Sci. Rep.* **9**, 14639 (2019).
29. Lunter, D. et al. Novel aspects of Raman spectroscopy in skin research. *Exp. Dermatol.* **31**, 1311–1329 (2022).
30. Paraskevaidi, M. et al. Clinical applications of infrared and Raman spectroscopy in the fields of cancer and infectious diseases. *Appl. Spectrosc. Rev.* **56**, 804–868 (2021).
31. Windom, B. C. & Hahn, D. W. Raman spectroscopy. in *Encyclopedia of Tribology* (eds. Wang, Q. J. & Chung, Y.-W.) 2742–2747 (Springer US, 2013). https://doi.org/10.1007/978-0-387-92897-5_1216.
32. Talari, A. C. S., Movasaghi, Z., Rehman, S. & Rehman, I. Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **50**, 46–111 (2015).
33. Kuhar, N., Sil, S., Verma, T. & Umapathy, S. Challenges in application of Raman spectroscopy to biology and materials. *RSC Adv.* **8**, 25888–25908 (2018).
34. Hernández, B., Pflüger, F., Kruglik, S. G. & Ghomi, M. Characteristic Raman lines of phenylalanine analyzed by a multiconformational approach. *J. Raman Spectrosc.* **44**, 827–833 (2013).
35. Aguiar, R. P., Falcão, E. T., Pasqualucci, C. A. & Silveira, L. Use of Raman spectroscopy to evaluate the biochemical composition of normal and tumoral human brain tissues for diagnosis. *Lasers Med. Sci.* **37**, 121–133 (2022).
36. Takeuchi, H. Raman structural markers of tryptophan and histidine side chains in proteins. *Biopolymers* **72**, 305–317 (2003).
37. Picot, F. et al. Data consistency and classification model transferability across biomedical Raman spectroscopy systems. *Transl. Biophotonics* **3**, e202000019 (2021).
38. Lemoine, É. 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* **144**, 6517–6532 (2019).
39. Rygula, A. et al. Raman spectroscopy of proteins: A review. *J. Raman Spectrosc.* **44**, 1061–1076 (2013).
40. Zhu, G., Zhu, X., Fan, Q. & Wan, X. Raman spectra of amino acids and their aqueous solutions. *Spectrochim. Acta A* **78**, 1187–1195 (2011).
41. Movasaghi, Z., Rehman, S. & Rehman, I. U. Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **42**, 493–541 (2007).
42. Czamara, K. et al. Raman spectroscopy of lipids: A review. *J. Raman Spectrosc.* **46**, 4–20 (2015).
43. Gerritsen, J. K. W. et al. Safe surgery for glioblastoma: Recent advances and modern challenges. *Neurooncol. Pract.* **9**, 364–379 (2022).
44. Hai, Y. et al. IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* **360**, 765–773 (2024).
45. Pisapia, D. J. Frozen Sections in Neuropathology. in *Frozen Section Pathology: Diagnostic Challenges* (eds. Borczuk, A. C., et al.) 247–264 (Springer International Publishing, Cham, 2021). https://doi.org/10.1007/978-3-030-71308-9_10.
46. Pinto, M. 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.* **24**, 1–10 (2019).
47. David, S. et al. *In situ* Raman spectroscopy and machine learning unveil biomolecular alterations in invasive breast cancer. *J Biomed. Opt.* **28**, 036009 (2023).
48. Aubertin, K. et al. Mesoscopic characterization of prostate cancer using Raman spectroscopy: Potential for diagnostics and therapeutics. *BJU Int.* **122**, 326–336 (2018).

49. Picot, F. et al. Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 1: Raman spectroscopy fiber-optics system and in situ tissue characterization. *J. Biomed. Opt.* **27**, 095003 (2022).
50. Grajales, D. et al. Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 2: In-vivo tumor-targeting using a classification model combining spectral and MRI-radiomics features. *J. Biomed. Opt.* **27**, 095004 (2022).
51. Leblond, F. et al. Subsecond lung cancer detection within a heterogeneous background of normal and benign tissue using single-point Raman spectroscopy. *J. Biomed. Opt.* **28**, 090501 (2023).
52. David, S. et al. Multispectral label-free Raman spectroscopy can detect ovarian and endometrial cancer with high accuracy. *J Biophotonics* **15**, e202100198 (2022).
53. Madsen, S. J. & Wilson, B. C. Optical properties of brain tissue. in *Optical Methods and Instrumentation in Brain Imaging and Therapy* (ed. Madsen, S. J.) 1–22 (Springer New York, 2013). https://doi.org/10.1007/978-1-4614-4978-2_1.
54. Akbarzadeh, A. et al. Experimental validation of a spectroscopic Monte Carlo light transport simulation technique and Raman scattering depth sensing analysis in biological tissue. *J. Biomed. Opt.* **25**, 105002 (2020).
55. Pedregosa, F. et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
56. Sheehy, G. et al. Open-sourced Raman spectroscopy data processing package implementing a baseline removal algorithm validated from multiple datasets acquired in human tissue and biofluids. *J. Biomed. Opt.* **28**, 025002 (2023).
57. Kopec, M., Błaszczyk, M., Radek, M. & Abramczyk, H. Raman imaging and statistical methods for analysis various type of human brain tumors and breast cancers. *Spectrochim. Acta A* **262**, 120091 (2021).
58. Grosset, A.-A. et al. Identification of intraductal carcinoma of the prostate on tissue specimens using Raman micro-spectroscopy: A diagnostic accuracy case–control study with multicohort validation. *PLoS Med.* **17**, e1003281 (2020).

Acknowledgements

This work is supported by the Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Collaborative Health Research Program (Canadian Institutes of Health Research and NSERC), and the TransMedTech Institute.

Author contributions

François Daoust: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Visualization, Writing – Original Draft, Writing – Review and Editing. Frédérick Dallaire: Methodology, Formal analysis, Visualization, Writing – Original Draft, Writing – Review and Editing. Hugo Tavera: Conceptualization, Methodology, Data Curation, Software, Visualization, Writing – Original Draft. Katherine Ember: Conceptualization, Writing – Original Draft, Writing – Review and Editing. Marie-Christine Guiot, Methodology, Investigation, Formal analysis, Resources. Kevin Petrecca: Methodology, Investigation, Resources, Supervision, Project administration, Funding acquisition, Writing – Review and Editing. Frédéric Leblond: Methodology, Supervision, Project administration, Funding acquisition, Visualization, Writing – Original Draft, Writing – Review & Editing.

Declarations

Competing interests

François Daoust, Kevin Petrecca and Frédéric Leblond are involved in Reveal Surgical, a medical device company that seeks to commercialize a Raman spectroscopy system for intraoperative cancer detection. The other authors of this manuscript do not have any competing interests in relation to the work described in this manuscript.

Human participants consent and ethics

Informed and written consent was obtained from all participants. The studies were approved by the Ethics Committees of the Montreal Neurological Institute and Hospital. All methods were performed in accordance with the relevant local and national guidelines and regulations.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-87109-1.

Correspondence and requests for materials should be addressed to F.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.