Accepted Manuscript

Spectral classification for diagnosis involving numerous pathologies in a complex clinical setting: A neuro-oncology example

Danielle Bury, Camilo L.M. Morais, Maria Paraskevaidi, Katherine M. Ashton, Timothy P. Dawson, Francis L. Martin

PII:	S1386-1425(18)30742-X
DOI:	doi:10.1016/j.saa.2018.07.078
Reference:	SAA 16347
To appear in:	*Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*
Received date:	23 May 2018
Revised date:	26 July 2018
Accepted date:	27 July 2018

Please cite this article as: Danielle Bury, Camilo L.M. Morais, Maria Paraskevaidi, Katherine M. Ashton, Timothy P. Dawson, Francis L. Martin , Spectral classification for diagnosis involving numerous pathologies in a complex clinical setting: A neuro-oncology example. Saa (2018), doi:10.1016/j.saa.2018.07.078

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Spectral classification for diagnosis involving numerous pathologies in a complex clinical setting: a neuro-oncology example

Danielle Bury[a], Camilo L M Morais[a], Maria Paraskevaidi[a], Katherine M Ashton[b], Timothy P Dawson[b], Francis L Martin[a]

[a]*School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK*

[b]*Department of Neuropathology, Royal Preston Hospital, Lancashire Teaching Hospitals NHS Trust, Preston PR2 9HT, UK*

Corresponding Author: Professor Francis L Martin, School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK; Tel.: +44(0)1772 896482; Email: flmartin@uclan.ac.uk

Conflicts of interest There are no conflicts to declare.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

Much effort is currently being placed into developing new blood tests for cancer diagnosis in the hope of moving cancer diagnosis earlier and by less invasive means than current techniques, *e.g.*, biopsy. Current methods are expected to diagnose and begin treatment of cancer within 62 days of patient presentation, though due to high volume and pressures within the NHS in the UK any technique that can reduce time to diagnosis would allow reduction in the time to treat for patients. The use of vibrational spectroscopy, notably infrared (IR) spectroscopy, has been under investigation for many years with varying success. This technique holds promise as is would combine a generally well accepted test (a blood test) with analysis that is reagent free and cheap to run. It has been demonstrated that, when asked simple clinical questions (*i.e.*, cancer *vs.* no cancer), results from spectroscopic studies are promising. However, in order to become a clinically useful tool, it is important that the test differentiates a variety of cancer types from healthy patients. This study has analysed plasma samples with attenuated total reflection Fourier-transform IR spectroscopy (ATR-FTIR), to establish if the technique is able to distinguish normal from primary or metastatic brain tumours. We have shown that when asked specific questions, *i.e.*, high-grade glioma *vs.* low-grade glioma, the results show a significantly high accuracy (100%). Crucially, when combined with meningiomas and metastatic lesions, the accuracy remains high (88-100%) with only minimal overlap between the two metastatic adenocarcinoma groups. Therefore in a clinical setting, this novel technique demonstrates potential benefit when used in conjuction with existing diagnostic methods.

Keywords: ATR-FTIR spectroscopy, biofluids, brain tumours, classification, sensitivity, specificity

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Introduction

Blood testing for cancer diagnostics is a popular ideal. It uses an acceptable patient test, *i.e.*, a blood test, which is minimally-invasive and machine analysed, can be run on a mass scale, and can target specific markers circulating in the blood. Whilst some cancers can be identified by the use of biomarkers, there are currently no such markers for primary or metastatic brain tumours, nor are any biomarkers yet involved in a mass-screening programme [1]. Brain tumours, both primary and metastatic, often present with a range of non-specific symptoms. The diagnostic process involves a combination of history taking, examination and radiology to determine the presence of a tumour and its possible origin [2]. There are specific radiological appearances that can help differentiate between primary and metastatic brain tumours; however, these rules do not always hold true [3]. A brain tumour may also be the first presentation of a metastatic cancer from elsewhere within the body; this accounts for up to a quarter of brain tumours [4]. Currently, a combination of radiological imaging and histology is used to detect the primary origin of a brain tumour. When metastatic, pathologists can apply immunohistochemical stains to formalin-fixed paraffin-embedded (FFPE) tissue, within which a combination of positive and negative stains can help determine a primary site of origin.

Over recent years the potential of vibrational spectroscopy has been touted as an 'inexpensive, high throughput and reagent-free' cancer diagnostic tool. *In vivo* studies have shown great promise using both tissue and blood component analysis with detection of cancer *vs.* non-cancer in many pilot studies showing promising results [5,6]. When considering biofluids, predominantly serum has been analysed for brain cancer, using attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy, with varying sample methods used [7]. This is due to the erythrocyte

ACCEPTED MANUSCRIPT

component in whole blood providing a strong interfering spectroscopic signal, likely masking the underlying changes seen in cancer *vs.* non-cancer patients [8]. The main limitations of these studies focus around different methods of sample preparation and analysis. No universal method of spectral analysis has yet been agreed. Butler *et al.* attempted to outline suggested methods of applying and analysing spectroscopic techniques on a variety of tissues, though it is unclear how widely accepted this has been [9].

Some studies have also compared the use of plasma or serum, with similar results. FTIR or Raman spectroscopy methods with biofluids have been used with a wide range of cancers; from head and neck (75% classification accuracy), bladder cancer (up to 80% accuracy), ovarian and endometrial cancer (96% and 81% respectively) and brain tumours (93% high-grade, 96% low-grade) [8, 10-16]. Spectral differences between gastric cancer patients *vs.* normal controls, leukaemia *vs.* normal controls and lung cancer *vs.* normal controls have all been demonstrated by a Chinese group [17-21]. These studies all demonstrate different points within the spectra of normal *vs.* malignant that exhibit specific changes. Whilst sensitivities and specificities have been high within the majority of these studies, they have all asked specific questions of the data, *e.g.*, cancer type *vs.* normal control, gliomas *vs.* meningioma.

There is then the limit of what question does ATR-FTIR need to answer. Within clinical medicine detection of a specific cancer is beneficial if suspected based on clinical suspicion. However, more frequently as symptoms are vague, a tool would be required to detect one from a number of primary cancers to provide early diagnostic utility. However, the majority of laboratory studies have focused on one cancer type and detecting its presence. It has been proposed that spectroscopy, more

ACCEPTED MANUSCRIPT

specifically, FTIR spectroscopy, could be used as a stand-alone tool for diagnostics or within screening [20]. However, when either a healthy person (*i.e.*, one without knowingly having the disease being tested for) presents for screening, or an unhealthy patient (*i.e.*, a person with a collection of symptoms not specifically demonstrating a defined underlying illness) with a non-specific history presents, the ability of a spectroscopic blood test to detect a cancer and its tissue origin from a range of options remains to be tested. Studies demonstrating detection of low-grade malignancies are also lacking; within brain tumours, the differentiation of low- and high-grade gliomas has been demonstrated to good effect [14]. Hands *et al.*, have also compared primary to metastatic tumours; however, this classed all brain tumours together *vs.* all types of metastasis [14]. On tissue, Krafft *et al.* have attempted to differentiate primary site from metastasis with limited effect. This study demonstrated that it was possible to separate by tumour type; however, adenocarcinomas for example, from two different primary sites gave markedly similar results [21]. Therefore in order to be clinically effective and aid other current diagnostic tools by adding value (*i.e.*, providing information by less invasive means or more quickly than current testing), any new test must be able to detect cancer, with high accuracy, and aim to localise the cancer from a range of primary sites, as is currently possible following biopsy with the use of immunohistochemistry.

The changes found within ATR-FTIR studies of peripheral blood components between cancer and non-cancer states are not yet clearly understood. Changes are seen at several wavelengths, *e.g.*, Apolipoprotein A1 within endometrial cancer [15] and several reports of changes in vibrations from C-O, C=O and C-H bonds associated with lipids and proteins within patients with brain metastases [14]. It has also been shown to be possible to detect epidermal growth factor receptor (EGFR)

ACCEPTED MANUSCRIPT

mutation status from peripheral blood of lung cancer patients [22]. This may be due to circulating tumour cells within the blood or in the case of ATR-FTIR spectroscopy-detected protein changes within the peripheral blood. These changes are likely caused by the release of enzymes and cytokines from the tumour as it grows.

Worldwide, neurological disorders account for >20% of disease burden and are frequently seen in General Practitioner (GP) clinics [23]. If a typical secondary care neurology clinic is used as an example, patients are referred *via* their GP due to underlying symptoms such as headaches. From the history, examination and any relevant investigations the clinician determines the underlying cause. MacDonald *et al.* looked at an unselected population from GP services in London who had been referred to tertiary neurology services. This was to understand the range of underlying diagnoses. They found non-tumour related pathology made up the vast majority of these presentations [24]. Similar results have been reported elsewhere, including a Nigerian study [23]. Hence any new investigative tool would need a high accuracy in picking out individual patients with a brain tumour against a backdrop of many confounding factors.

Therefore this study was performed to determine the accuracy of ATR-FTIR spectroscopy in detecting biochemical signatures of a range of primary and metastatic brain tumours from blood plasma when analysed in the same grouping. This approach aims to investigate whether ATR-FTIR spectroscopy is a useful tool to detect tumours within blood plasma when asked to differentiate on a wider-scale more akin to a typical clinical setting. Tumours were selected to encompass primary brain tumours and metastasis that commonly occur within the brain. Lung adenocarcinoma and malignant melanoma cause brain metastasis frequently, with lung adenocarcinoma resulting in the majority of brain metastasis [25-27]. Colorectal adenocarcinoma was

ACCEPTED MANUSCRIPT

chosen to compare two adenocarcinomas with often similar morphology but slightly different immunohistochemical profiles to determine if ATR-FTIR spectroscopy would be able to differentiate two adenocarcinomas of different primary origins. To the authors' knowledge this is the first paper to compare multiple tumour types using ATR-FTIR spectroscopy on plasma.

Methods

Plasma from 50 patients comprising normal, *i.e.*, no known brain tumour (*n*=10), glioma high-grade (*n*=5) or low-grade (*n*=5), meningioma (*n*=10) and brain metastasis patients, a mix of lung adenocarcinoma (*n*=7), colorectal adenocarcinoma (*n*=7) and malignant melanoma (*n*=6) patients were obtained from the Brain Tumour North West tissue bank (BTNW). This was under ethical approval number (RTB - ethics NRES14/EE/1270). These were stored at -80°C and defrosted prior to use. From the samples, 50 μL of plasma was pipetted onto a glass slide wrapped in aluminium foil. This has previously been shown to be as effective as slides such as CaF_2-coated windows [28].

The slides were left to dry overnight prior to spectral acquisition. ATR-FTIR spectra were collected using a Bruker TENSOR 27 FTIR spectrometer with Helios ATR attachment containing a diamond crystal internal reflective element and a 45^o incidence angle of IR beam. In the ATR module, the light goes through the sample in a reflectance mode and, after successive reflections between the sample and the ATR crystal, an evanescent wave is generated contained the attenuated IR signal of the sample [29-30]. For each case 32 scans with 8 cm^{-1} spectral resolution were taken at 10 randomly selected points. A new background spectrum was collected prior to each

ACCEPTED MANUSCRIPT

new sample, followed by cleaning of the crystal with distilled water. The sampling aperture was 250 μm × 250 μm and the mirror velocity was 2.2 Hz.

Computational analysis was then performed within a Matlab environment using IRootlab toolkit as a user interface [31]. Spectra were then pre-processed by cropping the region of interest (1850-800 cm^{-1}), followed by polynomial baseline correction and vector normalisation. Following this, principal component analysis-linear discriminant analysis (PCA-LDA) was performed to determine differences between the groups, along with PCA-linear discriminant classifier (PCA-LDC) to calculate the classification accuracy of each group. In addition, support vector machines (SVM) was employed for class differentiation using a radial basis function kernel. All PCA-LDC and SVM parameters (gamma value and number of support vectors) were optimized using leave-one-out cross-validation. Sub-dataset generation specification algorithm was employed for assigning 10% of the samples as external test dataset. The number of principal components was determined according to the maximum explained variance and lowest root mean square error of cross-validation. Statistical significance was then determined using a one-way ANOVA within PRISM (GraphPad Software) statistical analysis software.

Results

From the 50 cases 500 spectra (*i.e.*, 10 spectra per sample) were obtained. Following pre-processing (see Figure 1 for mean spectra), a PCA-LDA was performed to identify if the groups (or categories) are significantly different based upon their spectra, along with PCA-LDC to generate confusion matrices to look at the accuracy of the spectra in detecting each tumour type. Spectra appeared visibly different within the 1100-1000 cm^{-1} range.

ACCEPTED MANUSCRIPT

Analysis was then performed initially looking at normal *vs.* each tumour group and then combining all groups together to determine if they could be differentiated accurately from each other, as would occur in a typical clinical setting.

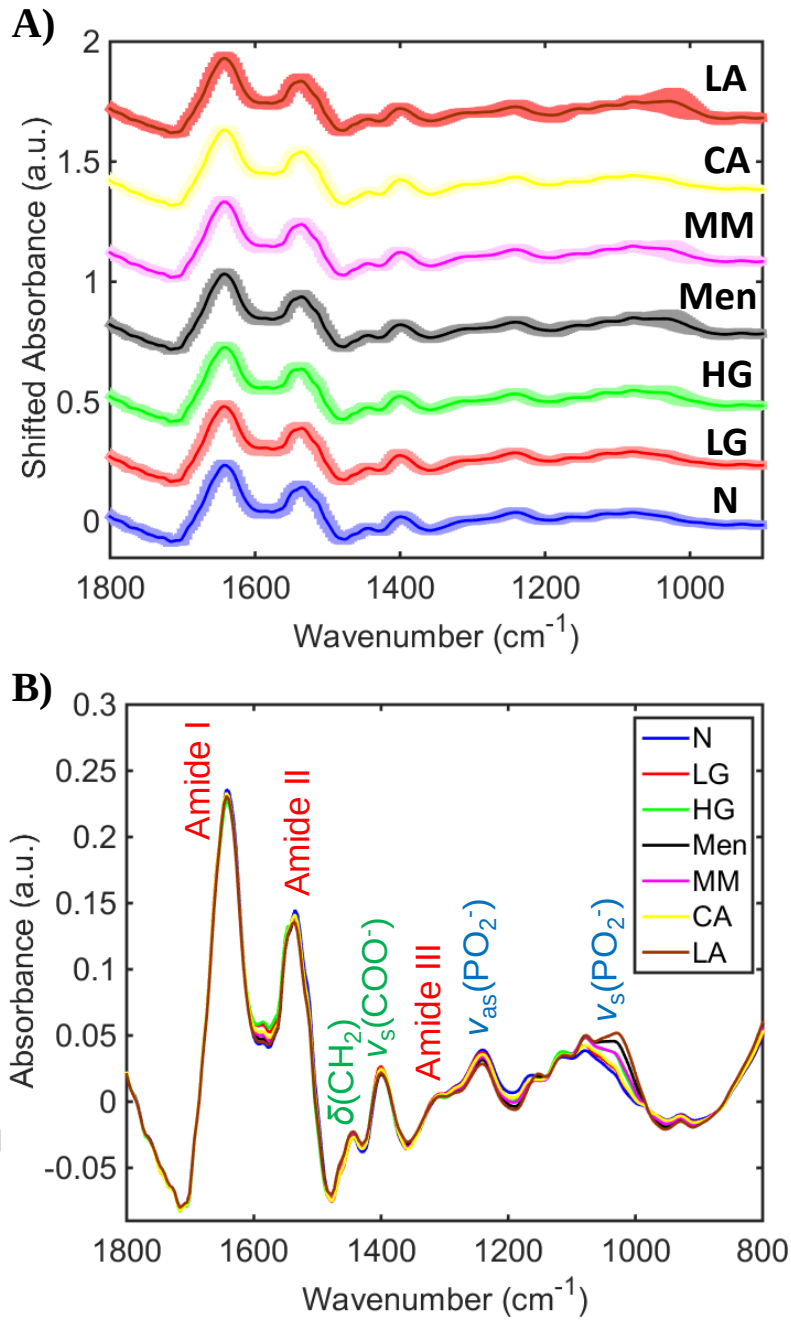


Figure 1. A) Mean pre-processed spectra with standard-deviation for all groups analysed; B) mean pre-processed spectra for all groups analysed with tentative assignment of main bands [29, 32]. Key: N, normal; LG, low-grade; HG, high-grade; Men, meningioma; MM, melanoma metastasis; CA, colorectal adenocarcinoma metastasis; LA, lung adenocarcinoma metastasis. Band assignment: ν_s, symmetric

ACCEPTED MANUSCRIPT

stretching; ν_{as}, asymmetric stretching; δ, deformation; blue: DNA/RNA absorptions; red: protein absorptions; green: lipids absorptions.

Figure 2 shows normal compared to low-grade and high-grade gliomas. It shows how well the spectra are separated based upon them being classed as normal or high-grade, particularly with SVM, with a100% accuracy (Table 1).

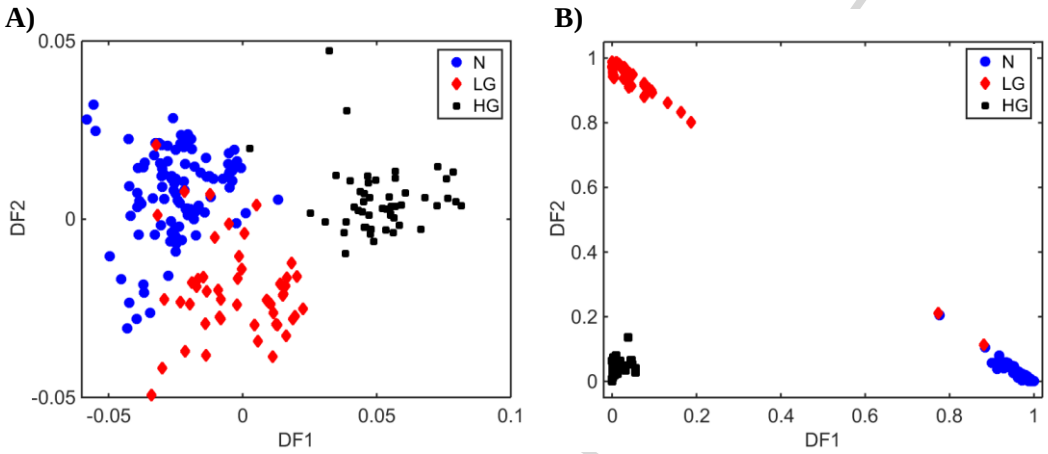


Figure 2. Discriminant function (DF) plot for PCA-LDC (A) and SVM (B) models. N stands for normal, LG for low-grade glioma, and HG for high-grade glioma samples. PCA-LDC performed with 10 principal components and SVM performed using 48 support vectors.

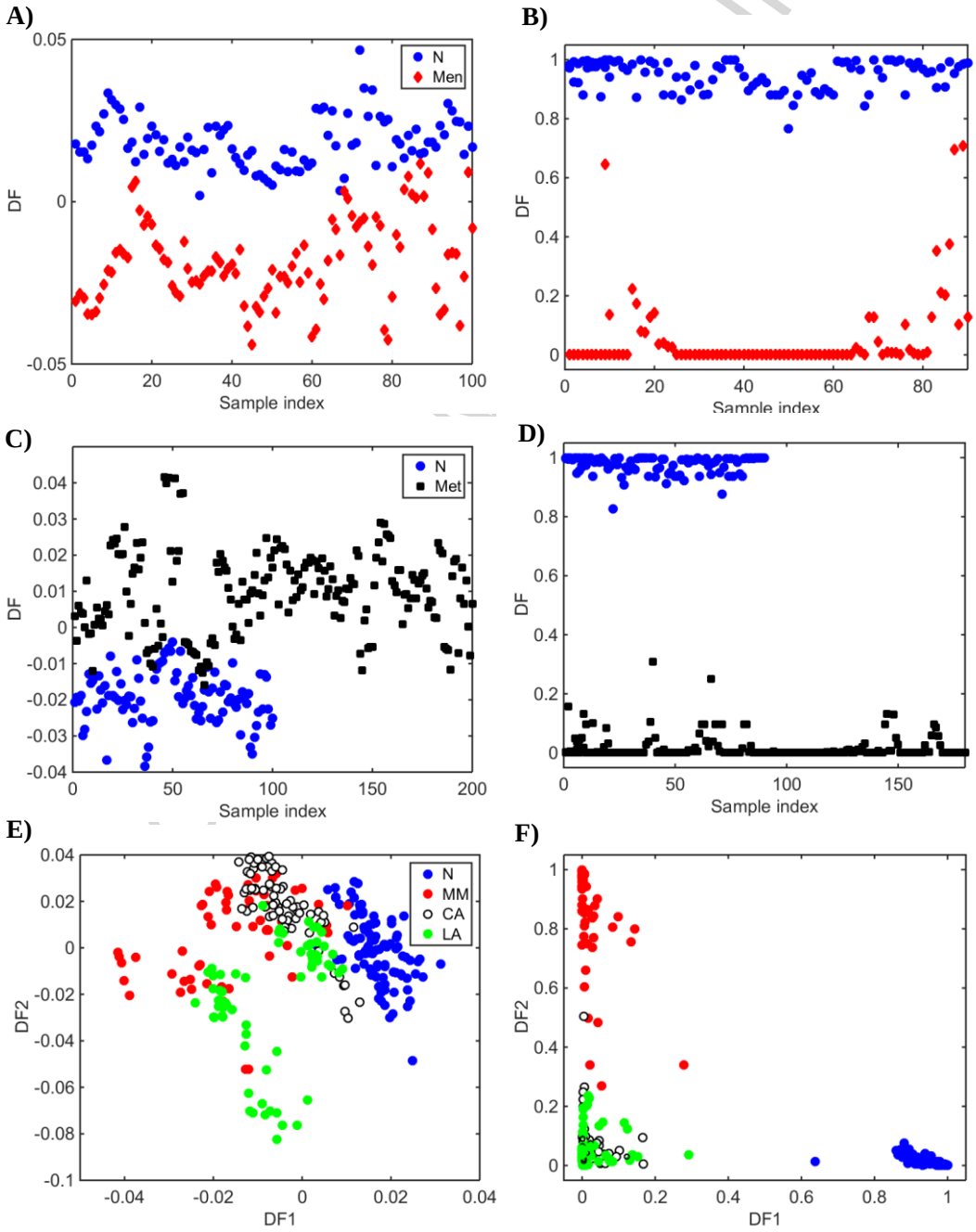
Table 1. Confusion table containing the accuracy for PCA-LDC and SVM models for distinguishing control (N), low-grade (LG) and high-grade (HG) gliomas.

	PCA-LDC (accuracy = 90%)			**SVM (accuracy = 100%)**		
	N	**LG**	**HG**	**N**	**LG**	**HG**
N	89%	11%	0%	100%	0%	0%
LG	16%	84%	0%	0%	100%	0%
HG	0%	2%	98%	0%	0%	100%

This differentiation between the three groups is statistically significant using a one-way ANOVA, as shown in Table S1 [see Supplementary Information (SI)]. Using the

ACCEPTED MANUSCRIPT

SVM method as the primary example, similar results are seen throughout, with minimal overlap seen within the two metastatic adenocarcinoma groups (Figures 3-6 and Tables 2-4). Given the morphological and immunohistochemical similarities between these two tumours this is to be expected. The PCA-LDC results, demonstrate high accuracy, which drops as the clinical question becomes increasingly complex (*i.e.*, by adding in different tumour types).



ACCEPTED MANUSCRIPT

Figure 3 Discriminant function (DF) plot for PCA-LDC and SVM models. (A) PCA-LDA and (B) SVM for normal (N) vs. meningioma (Men); (C) PCA-LDA and (D) SVM for normal (N) vs. metastasis (Met); (E) PCA-LDA and (F) SVM for normal (N) vs. different metastatic groups, colorectal adenocarcinoma (CA), lung adenocarcinoma (LA) and melanoma (MM). PCA-LDC performed with 10 principal components in (A)-(F), and SVM performed using 31 (A), 29 (D) and 140 (F) support vectors.

Table 2. Confusion table containing the accuracy for PCA-LDC and SVM models for distinguishing control (N), meningioma (Men), metastasis (Met), colorectal adenocarcinoma (CA), lung adenocarcinoma (LA), and melanoma (MM).

	PCA-LDC (accuracy = 87%)		**SVM (accuracy = 95%)**	
	N	**Men**	**N**	**Men**
N	89%	11%	91%	9%
Men	14%	86%	0%	100%

	PCA-LDC (accuracy = 94%)		**SVM (accuracy = 100%)**	
	N	**Met**	**N**	**Met**
N	96%	4%	100%	0%
Met	8%	92%	0%	100%

	PCA-LDC (accuracy = 63%)			
	N	**CA**	**LA**	**MM**
N	98%	1%	1%	0%
CA	19%	38%	22%	21%
LA	0%	16%	63%	21%
MM	3%	28%	15%	54%

	SVM (accuracy = 97%)			
	N	**CA**	**LA**	**MM**
N	100%	0%	0%	0%
CA	0%	88%	12%	0%
LA	0%	0%	100%	0%
MM	0%	0%	0%	100%

ACCEPTED MANUSCRIPT

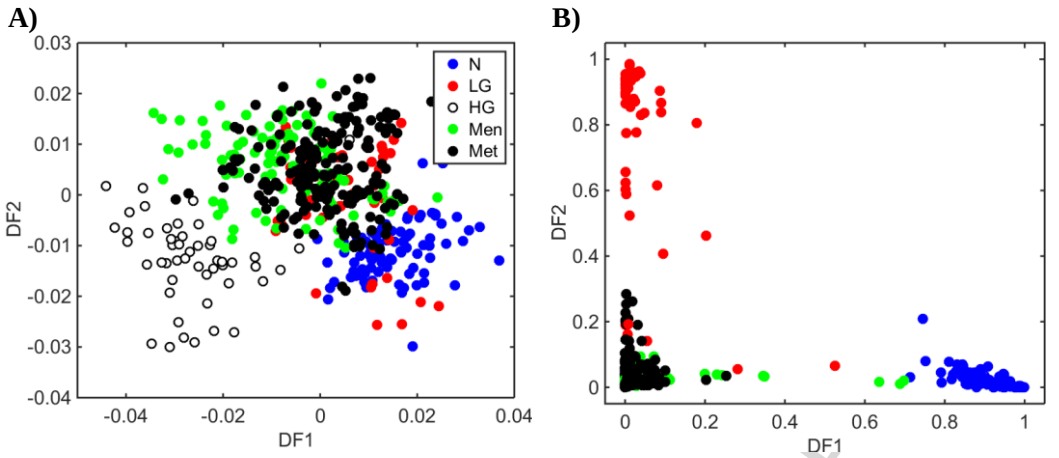


Figure 4 Discriminant function (DF) plot for (A) PCA-LDC and (B) SVM models. N stands for normal, LG for low-grade glioma, HG for high-grade glioma, Men for meningioma, and Met for metastasis samples. PCA-LDC performed with 10 principal components and SVM performed using 245 support vectors.

Table 3. Confusion table containing the accuracy for PCA-LDC and SVM models for distinguishing normal (N), low-grade glioma (LG), high-grade glioma (HG), meningioma (Men), and metastasis (Met).

	PCA-LDC (accuracy = 65%)					**SVM (accuracy = 97%)**				
	N	**HG**	**LG**	**Men**	**Met**	**N**	**HG**	**LG**	**Men**	**Met**
N	75%	0%	8%	10%	7%	91%	0%	0%	9%	0%
HG	10%	90%	0%	0%	0%	0%	100%	0%	0%	0%
LG	30%	1%	41%	10%	18%	0%	0%	100%	0%	0%
Men	4%	0%	3%	63%	30%	0%	0%	0%	100%	0%
Met	4%	0%	15%	23%	58%	0%	0%	5%	0%	95%

When examining the PCA-LDC, the classification accuracy drops further if the metastasis category is split by primary tumour location, with only the detection of high-grade glioma maintaining >90% (Figure 6). However, the use of SVM maintains

ACCEPTED MANUSCRIPT

accuracy approaching 100% for almost all groups. The overlap between meningioma and normal is difficult to explain without including patient factors within the algorithm, which is beyond the scope of this paper (Figure 6, Table 4).

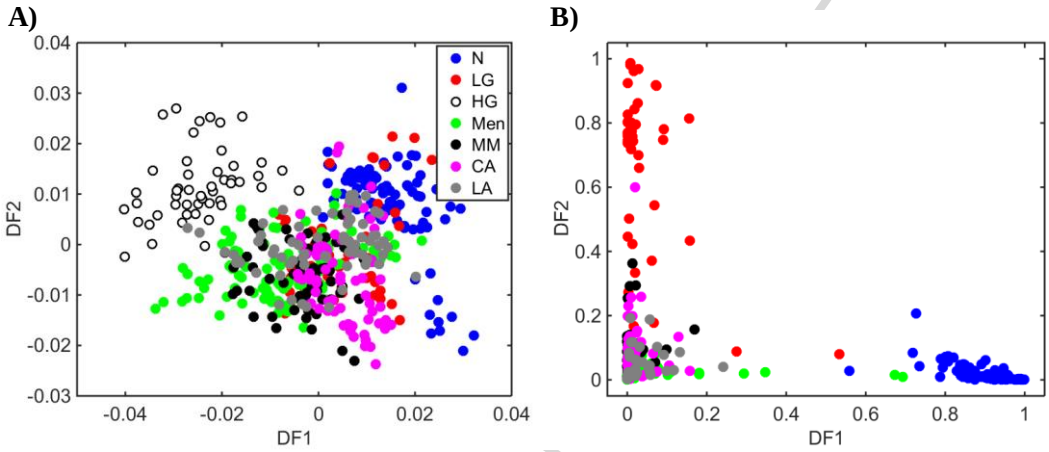


Figure 5 Discriminant function (DF) plot for (A) PCA-LDC and (B) SVM models. N stands for normal, LG for low-grade glioma, HG for high-grade glioma, Men for meningioma, MM for melanoma metastasis, CA for colorectal adenocarcinoma metastasis, and LA for lung adenocarcinoma metastasis. PCA-LDC performed with 10 principal components and SVM performed using 293 support vectors.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 4. Confusion table containing the accuracy for PCA-LDC and SVM models for distinguishing control (N), low-grade glioma (LG), high-grade glioma (HG), meningioma (Men), and melanoma metastasis (MM), colorectal adenocarcinoma metastasis (CA), and lung adenocarcinoma metastasis (LA).

	PCA-LDC (accuracy = 54%)							**SVM (accuracy = 97%)**						
	N	**LG**	**HG**	**Men**	**MM**	**CA**	**LA**	**N**	**LG**	**HG**	**Men**	**MM**	**CA**	**LA**
N	85%	6%	0%	0%	9%	0%	0%	91%	0%	0%	9%	0%	0%	0%
LG	25%	61%	4%	0%	10%	0%	0%	0%	100%	0%	0%	0%	0%	0%
HG	8%	0%	92%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%
Men	4%	0%	0%	43%	36%	3%	14%	0%	0%	0%	100%	0%	0%	0%
MM	0%	0%	0%	35%	22%	13%	30%	0%	0%	0%	0%	100%	0%	0%
CA	0%	3%	0%	5%	28%	50%	14%	0%	0%	0%	0%	0%	88%	12%
LA	18%	2%	0%	13%	29%	14%	24%	0%	0%	0%	0%	0%	0%	100%

Receiver operating characteristic (ROC) curves agree with the results in Tables 1-4, as can been seen in supplementary information Figures S2-S7 (see SI).

Discussion

With an ageing population, complex clinical presentations of cancer are becoming more frequent, with up to a quarter of all new brain tumour diagnoses representing metastases [4]. Therefore, this study was designed to mimic a clinical setting with a variety of primary and metastatic brain tumours. The aim was to determine if ATR-FTIR spectroscopy would be able to detect a brain tumour from peripheral blood and if so determine type. If this were to prove possible, the potential of ATR-FTIR in a frontline clinical setting as part of an initial workup of the patient

ACCEPTED MANUSCRIPT

may be demonstrable, as has previously been proposed [20]. The metastatic brain tumours were chosen based on both propensity for cranial spread and to provide tumours, which are morphologically similar (lung and colorectal adenocarcinomas) and a contrast (metastatic melanoma) to enable testing of ATR-FTIR spectroscopy to differentiate metastatic tumours by their primary origin [25-27]. Our results demonstrate that ATR-FTIR spectroscopy is able to detect patients with primary or metastatic brain tumours using plasma samples from peripheral blood. The SVM method of analysis, has demonstrated accuracy from 88-100% with minimal overlap between the two metastatic adenocarcinoma groups primarily, with ROC curves demonstrating specificity and $p < 0.001$. Given the similar morphological features of the two adenocarcinoma groups this was expected. Accuracy decreases slightly as the clinical question progresses from cancer *vs.* no cancer to which type of brain tumour and then primary metastatic location. With the differences seen within the spectra (Figure 1), this when combined with the classification accuracy, demonstrates the potential of ATR-FTIR spectroscopy to be built into clinical practise as a primary diagnostic step.

Previous studies performed by both this group and others have demonstrated that spectroscopic techniques can differentiate normal from cancer cases within biofluids alone. Owens *et al.* found a 93% accuracy when determining ovarian cancer patients from negative cases using plasma alone [11]. Gajjar *et al.* demonstrated similar findings when also comparing normal to ovarian and endometrial cancer patients [5]. Moving to bladder cancer, Ollesch *et al.* were also able to separate normal from bladder cancer patients [8]. Within brain tumours, Hands *et al.* have also shown the capabilities of ATR-FTIR spectroscopy with plasma, demonstrating classification of high and low grade glial tumours with accuracies of over 90%. These

ACCEPTED MANUSCRIPT

dip slightly when comparing to metastatic tumours [13-14]. Therefore our results, demonstrate comparable accuracies when compared across the published literature. They also take forward most of the studies above by combining different tumour types, both primary and metastatic.

Our study, has however, not included confounding patient factors within the analysis. The number or location of the brain tumours is unknown, nor with the metastatic patients the burden of metastatic disease. When patients initially present to a medical professional, it is unlikely they will have had all investigations completed and a known diagnosis, therefore for this initial study, it was felt to add other confounding factors into the algorithm would prove too complex and unlikely to yield usable results. Brain tumours account for a small number of diagnosed tumours per annum (11,432/359,960 in 2015) in comparison to extracranial tumours therefore it was felt limiting this study by size and location of brain tumours would be unlikely to be of benefit [33, 34].

To take this technique further forward and validate the results seen above clinically, it would be crucial to test a wide range of non-cancer disorders to ensure this would not affect the results. Current clinical trials for use of ATR-FTIR spectroscopy currently underway should be able to determine if within a clinical setting this technique can maintain the accuracies seen within this study. It would also be interesting to test cancer patients without brain metastases to see if this impacted the results and how this could be differentiated.

This study has shown that ATR-FTIR spectroscopy could play a role in plasma testing for both primary and extrinsic brain tumours. The advantages of ATR-FTIR spectroscopy in a clinical setting would be the relative low-cost of

ACCEPTED MANUSCRIPT

instrumentation; the fact that the technique is non-analyst dependent, thus not biased according to the analyst training; non-destructive, thus the sample could be reused for further analysis; and easily translated to automation, by using computer-based technologies. In addition, the use of ATR-FTIR spectroscopy clinically may be able to speed up the diagnostic process, though robust clinical studies would be required to eliminate any of the current diagnostic steps within the patient pathway.

As limitation, for such chemometric methods and instrumental techniques to be translated into a clinical setting, all possible experimental and computation configurations should be analysed, where all possible circumstances must be taken into consideration before clinical implementation. Evaluation of further biochemical factors and test with larger datasets are necessary before translation. This paper shows the potential usefulness of ATR-FTIR allied with chemometric methods as a biofluid screening tool, which remains to be further investigated.

Conclusions

This study demonstrates that ATR-FTIR spectroscopy is able to differentiate brain tumour types from blood plasma with accuracies of up to 100%. Further work is required to determine if this technique can discrimianat tumours from complex non cancer related disorders in order to all clinical medicine to harness the use of ATR-FTIR spectroscopy within the clincal pathway, hopefully decreasing time taken from clinical presentation to diagnosis.

Acknowledgements The authors would like to acknowledge the support from Rosemere Cancer Foundation, the Brain Tumour North West RTB and the Sidney Driscoll Neuroscience Foundation for their support. CLMM would like to thank CAPES-Brazil (grant 88881.128982/2016-01) for financial support. MP acknowledges Rosemere Cancer Foundation for funding. The authors would also like to thank Kestra Dawson for her help with production of a graphical abstract.

ACCEPTED MANUSCRIPT

References

1 W. D. Hazelton, E. G. Luebeck, Biomarker-Based Early Cancer Detection: Is it Achievable? Science Translational Medicine 3 (2011) 109-114.

2 Cancer Research UK. Brain Tumours, Tests to Diagnose. https://www.cancerresearchuk.org/about-cancer/brain-tumours/getting-diagnosed/tests-diagnose. Accessed July 2018.

3 D. Ellison, S. Love, L. Chimelli, B. N. Harding, J. S. Lowe, H. V. Vinters, S. Brandner, W. H. Yong. A reference text of CNS Pathology. Neuropathology. 3rd Edition, 2013. Elsevier Mosby.

4 L. Bekaert, E. Emery, G. Levallet, E. Lechapt-Zalcman, Histopathologic diagnosis of brain metastasis: current trends in management and future considerations. Brain Tumour Pathology 34 (2017) 8-19.

5 H. Byrne, M. Baranska, G. J. Pupples, N. Stone, B. Wood, K. M. Gough, P. Lasch, P. Heraud, J. Sule-Suso, G. Sockalingum, Spectropathology for the next generation: Quo vadis? Analyst 140 (2015) 2066-2073.

6 A. L. Mitchell, K. B. Gajjar, G. Theophilou, F. L. Martin, P. L. Martin-Hirsch, Vibrational Spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting. Journal of Biophotonics 7 (2014) 153-165.

7 C. Hughes, M. Brown, G. Clemens, A. Henderson, G. Monjardez, N. W. Clarke, P. Gardner, Assessing the challenges of Fourier transform infrared spectroscopic analysis of blood serum. Journal of Biophotonics 7 (2014) 180-188

8 K. Gajjar, J. Trevisan, G. Owens, P. J. Keating, N. J. Wood, H. F. Stringfellow, P. L. Martin-Hirsch, F. L. Martin, Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer. Analyst 138 (2013) 3917-3926.

9 H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardener, P. L. Martin-Hirsch, M. J. Walsh, M. R. McAinsh, N. Stone, F. L. Martin, Using Raman spectroscopy to characterize biological materials. Nature Protocols 11 (2016) 664-687.

10 K. Gajjar, L. D. Heppenstall, W. Pang, K. M. Ashton, J. Trevisan, I. I. Patel, V. Llabjani, H. F. Stringfellow, P. L. Martin-Hirsch, T. Dawson, F. L. Martin, Diagnostic segregation of human brain tumours using Fourier transform infrared and/or Raman spectroscopy coupled with discriminant analysis. Analytical Methods, 5 (2012) 89-102.

11 F. L. Martin, J. G. Kelly, V. Llabjani, P. L. Martin-Hirsch, I. I. Patel, J. Trevisan, N. J. Fullwood and M. J. Walsh, Distinguishing cell types or populations based on the computational analysis of their infrared spectra. Nature Protocols 5 (2010) 1748-1760.

12 J. Ollesch, M. Heinze, H. M. Heise, T. Behrens, T. Bruning, K. Gerwert, It's in your blood: spectral biomarker candidates for urinary bladder cancer from automated FTIR spectroscopy. Journal of Biophotonics 7 (2014) 210-221

13 J. R. Hands, K. M. Dorling, P. Abel, K. M. Ashton, A. Brodbelt, C. Davis, T. Dawson, M. D. Jenkinson, R. W. Lea, C. Walker, M. J. Baker, Attenuated total reflection fourier transform infrared (ATR-FTIR) spectral discrimination of brain tumour severity from serum samples. Journal of Biophotonics 7 (2014) 189-199

14 J. R. Hands, G. Clemens, R. Stables, K. Ashton, A. Brodbelt, C. Davis, T. P. Dawson, M. D. Jenkinson, R. W. Lea, C. Walker, M. J. Baker, Brain tumour

ACCEPTED MANUSCRIPT

differentitation: rapid stratified serum diagnostics via attenuated total refection Fourier-transform infrared spectroscopy. Journal of Neurooncology 127 (2016) 463-472

15 G. Owens, K. Gajjar, J. Trevisan, S. W. Fogarty, S. E. Taylor, B. Da Gama-Rose, P. L. Martin-Hirsch, F. L. Martin, Vibrational biospectroscopy coupled with multivariate analysis extracts potentially diagnostic features in blood plasma/serum of ovarian cancer patients. Journal of Biophotonics 7 (2014) 200-209

16 K. M. G. Lima, K. B. Gajjar, P. L. Martin-Hirsch, F. L. Martin, Segregation of ovarian cancer stage exploiting spectral biomarkers derived from blood plasma or serum analysis: ATR-FTIR spectroscopy coupled with variable selection methods. Biotechnology Progress 31 (2015) 832-839.

17 D. Sheng, Y. Wu, X. Wang, D. Huang, X. Chen, X. Liu, Comparison of serum from gastric cancer patients and from healthy persons using FTIR spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 116 (2013) 365-369.

18 D. Sheng, X. Liu, W. Li, Y. Wang, X. Chen, X. Wang, Distinction of leukemia patients' and healthy persons' serum using FTIR spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 101 (2013) 228-232.

19 X. Wang, X. Shen, D. Sheng, X. Chen, X. Liu. FTIR spectroscopic comparison of serum from lung cancer patients and healthy persons. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 122 (2014) 193-197.

20 C. Hughes, M. J. Baker, Can mid-infrared biomedical spectroscopy of cells, fluids and tissue aid improvements in cancer survival? A patient paradigm. Analyst 141 (2016) 467-475.

21 C. Krafft, L. Shapoval, S. B. Sobottka, K. D. Geiger, G. Schackert, R. Salzer, Identification of primary tumors of brain metastases by SIMCA classification of IR spectroscopic images. Biochimica et Biophysica Acta 1758 (2006) 883-891.

22 M. Brevet, M. L. Johnson, C. G. Azzoli, M. Ladanyi, Detection of EGFR mutations in plasma DNA from lung cancer patients by mass spectrometry genotyping is predictive of tumor EGFR status and response to EGFR inhibitors, Lung Cancer 73 (2011) 96-102.

23 I. O. Onwuekwe, B. Ezeala-Adikaibe, Prevalence and Distribution of Neurological Disease in a Neurology Clinic in Enugu, Nigeria. Ann Med Health Sci Res, 1, (2011) 63-67.

24 B. K. MacDonald, O. C. Cockerell, J. W. A. S. Sander, S. D. Shorvon, The incidence and lifetime prevalence of neurological disorders in a prospective community-based study in the UK. Brain 123 (2000) 665-676.

25 F. G. Davis, T. A. Dolecek, B. J. McCarthy, J. L. Villiano, Toward determining the lifetime occurance of metastatic brain tumours estimated from 2007 United States cancer incidence data. Neuro-Oncology 14 (2012) 1171-1177.

26 Q. Huang, X. Huang, Predictive biochemical-markers for the development of brain metastases from lung cancer: clinical evidence and future directions, Cancer Epidemiology 37 (2013) 703-707.

27 J. J. Renfrow, G. J. Lesser, Molecular subtyping of brain metastases and implications for therapy. Current Treatment Options in Oncology 14 (2013) 514-527.

ACCEPTED MANUSCRIPT

28 L. Cui, H. J. Butler, P. L. Martin-Hirsch, F. L. Martin, Aluminium foil as a potential substrate for ATR-FTIR, transflection FTIR or Raman spectrochemical analysis of biological specimens. Analytical Methods 8, (2016) 481-487.

29 M.J. Baker, J. Trevisan, P. Bassan, R. Bharagva, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys, C. Hughes, P. Lasch, P. L. Martin-Hirsch, B. Obinaju, G. D. Socklingum, J. Sule-Suso, R. J. Strong, M. J. Walsh, B. R. Wood, P. Gardner, F. L. Martin. Using Fourier transform IR spectroscopy to analyze biological materials. Nature Protocols 9 (2014) 1771-1791.

30 F. Zaera. New advances in the use of infrared absorption spectroscopy for the characterization of heterogeneous catalytic reactions. Chemical Society Reviews 43 (2014) 7624-7663.

31 J. Trevisan, P. P. Angelov, A. D. Scott, P. L. Carmichael, F. L. Martin, IRootLab: a free and open-source MATLAB toolbox for vibrational biospectroscopy data analysis. Bioinformatics 29 (2013) 1095-1097.

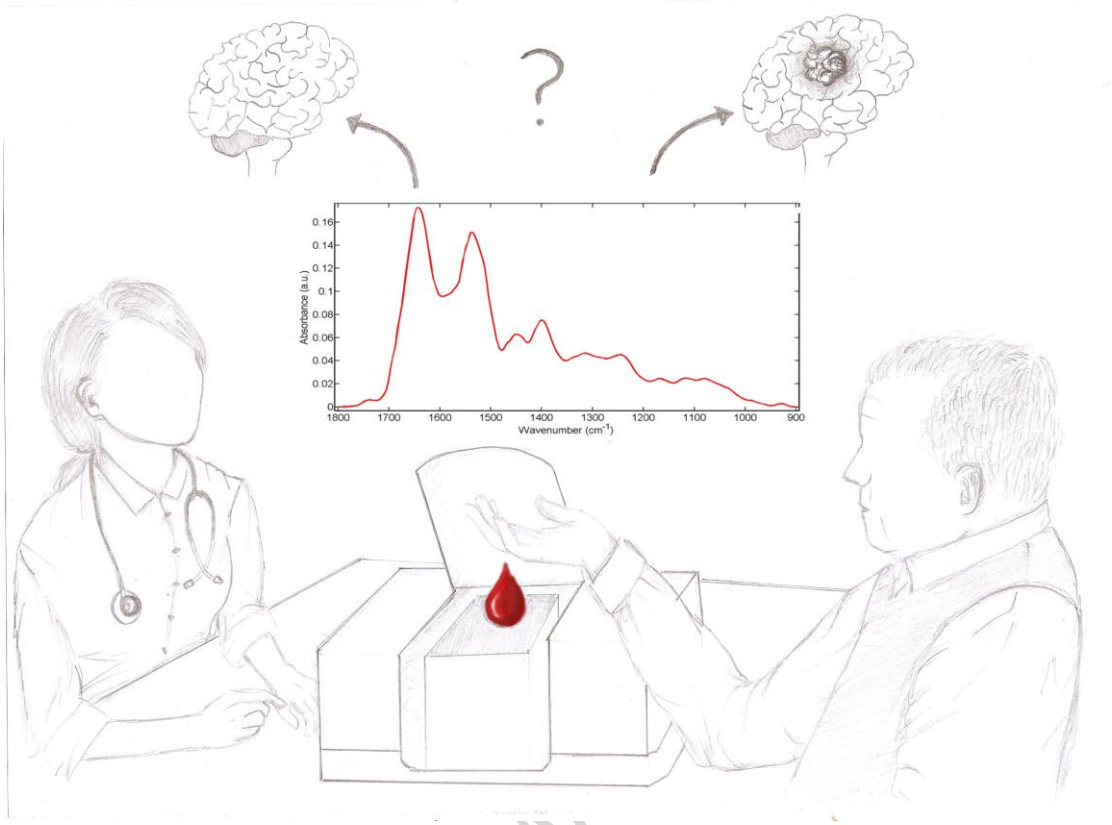
32 Z. Movasaghi, S. Rehman, I. Rehman. Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues. Applied Spectroscopy Reviews 43 (2008) 134-179.

33 Cancer Research UK. Brain and other CNS and intracranial tumour cancer statistics. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/brain-other-cns-and-intracranial-tumours#heading-Two. Accessed July 2018.

34 Cancer Research UK. Cancer statistics for the UK. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk. Accessed July 2018.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT



Graphical abstract

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Highlights

- Vibrational spectroscopy can diagnose brain tumours in biofluids
- Reagent-free and inexpensive blood test
- Can be applied in a clinical setting with multiple tumour types
- Accuracy in terms of sensitivity / specificity remains high
- ATR-FTIR spectroscopy could be a robust diagnostic technique

ACCEPTED MANUSCRIPT