

Mini-review

Biofluid Diagnostics by FTIR Spectroscopy: A Platform Technology for Cancer Detection.

Alexandra Sala^a, David J. Anderson^a, Paul M. Brennan^b, Holly J. Butler^c, James M. Cameron^a, Michael D. Jenkinson^d, Christopher Rinaldi^a, Ashton G. Theakstone^a and Matthew J. Baker^{a,c,*}

^aWestCHEM, Department of Pure and Applied Chemistry, University of Strathclyde, Technology and Innovation Centre, 99 George Street, Glasgow, G1 1RD, UK

^bTranslational Neurosurgery, Department of Clinical Neurosciences, Western General Hospital, Edinburgh, EH4 2XU, UK

^cClinSpec Diagnostics Ltd., Technology and Innovation Centre, 99 George Street, Glasgow, G1 1RD, UK

^dInstitute of Translational Medicine, University of Liverpool & The Walton Centre NHS Foundation Trust, Lower Lane, Fazakerley, Liverpool, L9 7IJ, UK

ARTICLE INFO

Keywords: Infrared Spectroscopy, ATR-FTIR, Cancer, Early Detection, Biofluids, Serum, Brain Tumour

ABSTRACT

Fourier Transform Infrared Spectroscopy (FTIR) has been largely employed by scientific researchers to improve diagnosis and treatment of cancer, using various biofluids and tissues. The technology has proved to be easy to use, rapid and cost-effective for analysis on human blood serum to discriminate between cancer *versus* healthy control samples. The high sensitivity and specificity achievable during samples classification aided by machine learning algorithms, offers an opportunity to transform cancer referral pathways, as it has been demonstrated in a unique and recent prospective clinical validation study on brain tumours. We herein highlight the importance of early detection in cancer research using FTIR, discussing the technique, the suitability of serum for analysis and previous studies, with special focus on pre-clinical factors and clinical translation requirements and development.

1. Introduction

1.1 Early Diagnosis of Cancer

Cancer is a significant and growing global health burden. It is the leading cause of mortality in people aged younger than 70 years old in 48 countries, and the second leading cause in another 43 countries, according to the estimates of premature deaths from the World Health Organization (WHO) in 2015 [1]. Cancer incidence and mortality are rapidly growing worldwide, as consequence of population ageing and growth; the increasing rates reflect also the changes in cancer prevalence and distribution of important risk factors for cancer, such as diet and smoking [1]. In 2018, an estimated 9.6 million people died of cancer worldwide, predicted to rise to 13.0 million by 2030 [2]. In the twenty-first century, cancer is forecast to overtake ischaemic heart disease as the leading cause of mortality, and be the most significant obstacle to increase life expectancy [1,3].

Early diagnosis of cancer is a key strategy to reduce mortality rates, as earlier treatment is more effective. Furthermore, continuous progress in molecular biology and cancer development studies has resulted in more targeted and effective therapies for many cancers [4].

Screening is an important strategy for early detection of several cancers and is based on the identification of tumour-specific biomarkers, found in multiple biofluids or body tissues. These same biomarkers may be used for different purposes, including screening, diagnosis, staging or disease monitoring [5].

The accuracy of screening tests represents the principal factor in the evaluation of clinical utility. Sensitivity and specificity are the statistical parameters taken into consideration during such

evaluation; the former is related to the ability of the test to identify true positives over false negatives and the latter to identify true negatives over false positives [5]. Prostate-Specific Antigen (PSA, prostate tumour marker) and other cancer antigen tests such as CA 19-9 (gastrointestinal and pancreatic tumour marker) and CA 125 (ovarian tumour marker) are currently used in cancer detection. They have high specificity values around and over 90%, however their sensitivity ranges from 50-62% (CA 125) [6] to 72% (PSA) [7] and 79-81% (CA 19-9) [8]. When evaluating the efficacy of these current biomarkers in cancer screening, their sensitivity and specificity in the context of disease prevalence yields a resultant low Positive Predictive Value (PPV). PPV is the likelihood that a positive test result predicts a true positive cancer diagnosis; therefore, the use of screening tests with low PPVs can lead to over-diagnosis and over-treatment which are both unwarranted and unnecessary [5]. The majority of existing cancer biomarkers are unsuitable for use in national or population screening programmes, and they are currently used only on an *ad-hoc* basis in symptomatic individuals. Following a screening test, a patient with a positive/abnormal test result will follow a diagnostic pathway as shown in Figure 1. This highlights the importance of accurate and effective initial tests as a route into clinical investigation and treatment.

Many different scientific fields and techniques have been explored to improve early detection of cancer. Research has focused on identification of circulating tumour DNA (ctDNA), although this is yet to effectively penetrate clinical practice, not least because of technical challenges and cost [9]. By contrast, infrared (IR) spectroscopy, which can be employed for the analysis of multiple different biofluids, has been widely proven by numerous proof-of-principle studies [10–14]. It is low-cost, non-invasive, rapid, highly sensitive and specific – features that can prove it as a potentially powerful clinical tool, which could make early detection possible across many different cancers [15].

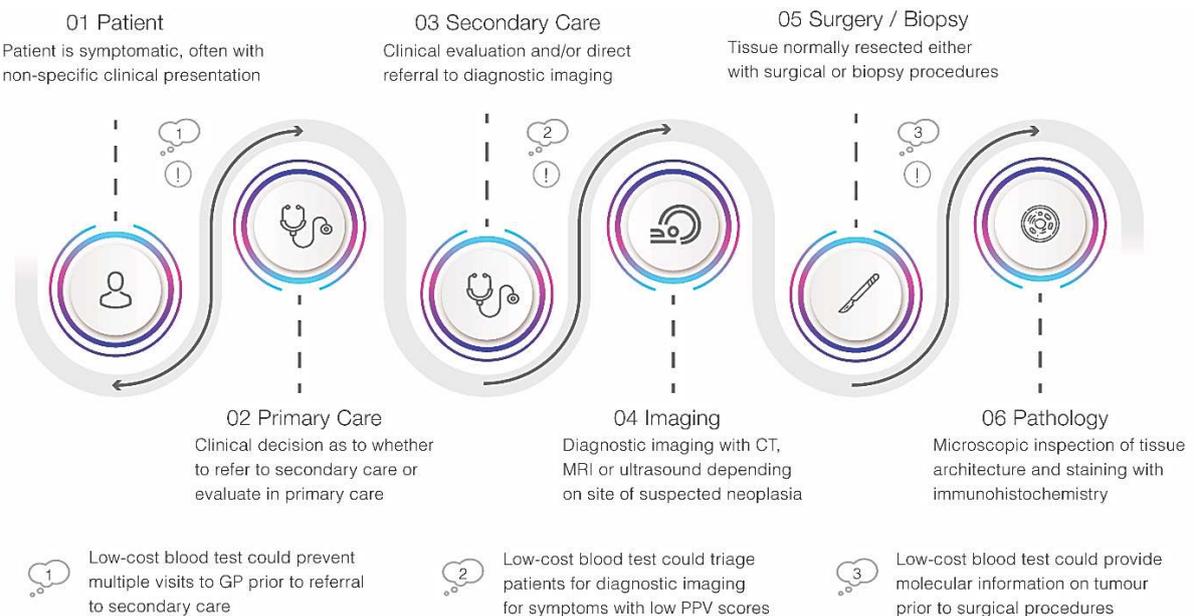


Figure 1 – Stepwise process of the current cancer detection path in clinical setting.

1.2 Infrared Spectroscopy and the Attenuated Total Reflection technique

Fourier Transform Infrared (FTIR) spectroscopy is a widely utilised analytical technique for the spectral characterisation of a wide range of solids, liquids and gases; including both chemical and biological species [16,17]. Traditional FTIR spectroscopy involves directly transmitting the infrared beam through the sample; however, a more recent and simpler approach involves an Attenuated Total Reflection (ATR) mode. This technique measures the changes in infrared radiation from an internally reflected beam when contact is made with the sample that is directly placed on top of the ATR crystal

[17]. Experiments with transmission mode showed that resultant spectra are strongly influenced by a range of physical phenomena, such as light scattering, refraction and dispersion. As ATR mode operates with different optical properties compared to transmission spectroscopy, it overcomes those spectral issues [18]. At a certain predetermined angle, the infrared radiation is directed onto a crystal with a high refractive index. Within the crystal, there is an internal reflectance which creates an evanescent wave that protrudes beyond the surface and into the sample placed on top. The sample absorbs energy within the infrared spectrum region dependent on its molecular composition, and these absorptions attenuate the IR beam in a detectable way. The attenuated energy beam is then passed to the detector by exiting the opposite end of the crystal [17].

There are two main requirements for the successful spectral acquisition utilising ATR mode: (i) the refractive index of the sample must be lower than that of the crystal for internal reflectance to occur, and (ii) the evanescent wave only protrudes beyond the crystal 0.5-5 μm therefore the sample must be in direct contact with the surface [17]. ATR crystals are available in zinc selenide (ZnSe), zinc sulphide (ZnS), germanium (Ge), silicon (Si) and diamond; however, ZnSe, Ge and diamond are the most commonly used due to their material properties [19]. Germanium is particularly useful for highly absorbing samples, such as rubbers, due to its high refractive index (4.01), however the Ge spectral region only ranges from 600-5,000 cm^{-1} . Zinc selenide has a wider spectral region (500-20,000 cm^{-1}), however it is only suitable for liquids or soft materials and cannot be used on samples with a pH less than 5 or greater than 9. Diamond is universally the preferred material despite the higher initial costs, as it has a greater spectral region (10-45,000 cm^{-1}), higher resistance and is chemically inert [19].

The use and applications of vibrational spectroscopic techniques for biological studies continues to increase as it is reproduceable, non-destructive, simple and only requires small amounts of sample material [20]. The FTIR peaks can be assigned to specific vibrations of chemical bonds or functional groups within the molecule, therefore providing information on the biochemical arrangement [21]; Figure 2 shows the typical spectrum of human blood serum analysed with Attenuated Total Reflection – Fourier Transform Infrared (ATR-FTIR) spectroscopy and the identifiable biomolecules [22]. Any molecular changes associated with disease can be optically probed and detected using FTIR spectroscopy as it is sensitive to the full range of biomolecular classes.

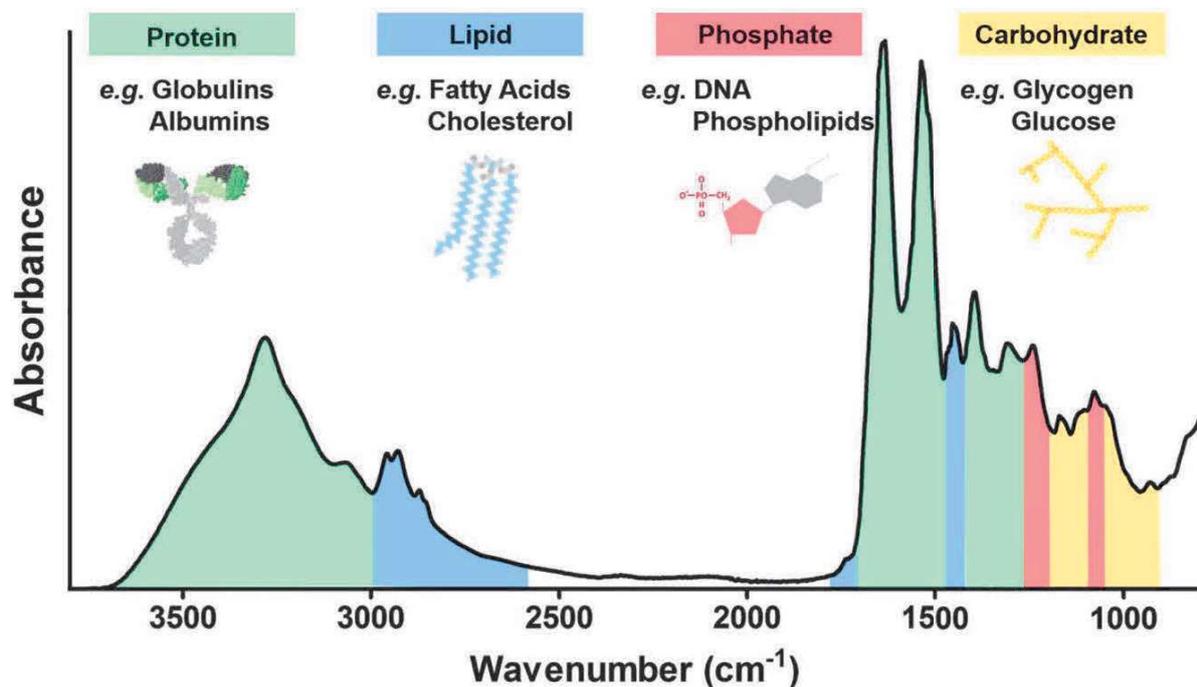


Figure 2 – Human blood serum spectrum acquired by ATR-FTIR spectroscopy with relative biomolecules associated to peaks. Adapted with permission from ref. [22]

1.3 Suitability of Serum

The majority of research in the biomedical vibrational spectroscopy field have been based on the spectroscopic analysis of human tissue, with many pilot studies suggesting it is possible to distinguish between healthy and cancerous tissue, as well as benign and malignant tumours [21]. Various malignancies - such as breast, lung, colon and prostate tissues - have previously been studied which has provided a platform of encouraging results [23–27]. Unfortunately, the technique has yet to make a successful transition into routine clinical practice, despite the abundance of promising publications [28].

In the past decade there has been further interest in biofluid spectroscopy [29]. This has been mainly due to the ease of sample collection and handling, and minimal preparation required prior to spectroscopic analysis [30]. ATR-FTIR spectroscopy is particularly useful for the analysis of biofluids, as only minute volumes (μL) of sample are required. With the recent advances in of ATR accessories [31] and in complex data analytics [32], high-throughput ATR-FTIR has the potential to deliver a patient's diagnosis within minutes, using a single blood sample.

Blood components, such as serum and plasma, are commonly analysed in clinical tests, as they carry information regarding intra- and extra-cellular events. Blood serum is the liquid part of the blood that contains no clotting factors or blood cells [33]. It exists as the most complex biofluid, containing over 20,000 different proteins, ranging from the larger abundant proteins like serum albumin (50 g L^{-1}) to the smaller protein molecules such as troponin (1 ng L^{-1}) [34]. Blood serum perfuses the organs throughout the entire human body, hence it gains proteomes from surrounding tissues and cells [35]. The low molecular weight fraction of serum, known as the peptidome, is thought to contain cancer-specific diagnostic information, making the spectroscopic biosignature of serum ideal for detecting disease states [36].

Serology is already commonly employed in the clinic, as serum is the preferred specimen for many diagnostic tests [37]. The concentrations of various biomarkers and diagnostic indicators can be measured in serum samples [38,39]. At present, the levels of biomolecular components such as albumin, bilirubin, creatinine and troponin – along with many others, including all the previously discussed cancer biomarkers (PSA, CA 19-9 and CA 125) – are all determined using serum tests in the health services worldwide. Therefore, an advantage of using serum for the triage of cancer is the potentially smooth transition into current clinical pathways. Since blood serum tests at the primary care level are already ordered as standard, the addition of serum spectroscopy into the clinical pathway would not significantly disrupt existing practices within the clinic.

1.4 ATR-FTIR Spectroscopy for Translation to the Clinic

ATR-FTIR is a promising analytical tool for cancer serum diagnostics for a range of cancers including brain [40], lung [41] and ovarian [42], however clinical translation remains elusive. The transition from research to routine clinical testing has faced several major technological challenges, as well as the challenges presented from the current patient pathway and the clinician acceptance of a new diagnostic test. Principally, the IRE is fundamental to spectroscopic interrogation of serum samples and must possess a greater refractive index than the biofluid at the analytical interface to facilitate production of evanescent waves. As previously mentioned, IREs comprise diamond, germanium or zinc selenide substrates, characteristic of high raw material and fabrication costs [31]; these ultimately limit the applicability of ATR-FTIR for high-throughput clinical testing. Similarly, the IRE is permanently secured to the top-plate of a spectrometer that restricts spectroscopic analysis to a step-wise approach where clinical biofluids must be individually spotted, dried, analysed and cleaned, which is impractical for high-volume and fast-paced clinical environments. Furthermore, the fixed nature of the IRE poses significant biofluid contamination risks, which makes it unsuitable for clinical testing where a disposable diagnostic platform is highly desirable.

Recently, the development of micro-fabricated silicon slides for use as IREs has demonstrated considerable promise to overcome current practical limitations of ATR-FTIR and facilitate clinical translation for cancer diagnostics. In particular, the low-cost, disposable test platform has demonstrated the ability to batch process serum samples and diagnose cancer versus non-cancer with a sensitivity and specificity of 93.2% and 92.8% respectively in a large, retrospective cohort of 724 patients [31]. Critical to commercial translation of the proposed diagnostic test has been the careful consideration of current challenges in the cancer diagnostic pathway and the best implementation of the technology for earlier diagnosis for patients, whilst proving economically viable for health care services. To this extent, health economic studies were conducted, showing that the technology would be cost-effective as a triage tool in primary or secondary care at test sensitivity and specificity of $\geq 80\%$ [22]. Consequently, the technology recently participated in a prospective clinical validation study at the Western General Hospital in Edinburgh (Scotland), where the test performed with sensitivity and specificity of 83.3% and 87.0% for the evaluation of symptomatic patients referred to secondary care through open access CT to eliminate brain tumour diagnosis [31]. Hence, ATR-FTIR represents a low-cost screening method for symptomatic patients in primary care, thereby streamlining the current imaging diagnostic pathway with substantial cost-saving implications for health services. Furthermore, the diagnostic platform highlights the clinical steps required before adoption of novel technologies can be considered, and the importance for emerging technologies to integrate and compliment current clinical pathways to facilitate future clinical translation of ATR-FTIR within health care settings.

2. Infrared Spectroscopic Analysis of Biofluids: Modern Attempt at Cancer Detection

Vibrational spectroscopy affords the opportunity to investigate the molecular composition of organic and inorganic compounds. However, only in the recent years, it has seen an increasing interest in the numerous potential applications of vibrational spectroscopy, particularly in the clinical environment. Although the gold standard for cancer diagnosis remains histology, many studies investigated the use of infrared (IR) spectroscopy through the analysis of cells, tissues and biofluids, as the simplicity, rapidity and cost-effectiveness of the technique may represent a step forward to translation into the clinic and early detection of cancer [43,44].

We discuss here the most relevant proof-of-concept studies of the last decade investigating cancer detection using high-throughput (HT-) and ATR-FTIR spectroscopy to analyse different biofluids, particularly human blood serum.

In 2010, Backhaus *et al.* [45] were one of the first to analyse serum with FTIR spectroscopy in 2010, in order to detect breast cancer. 98 patients aged 31-96 years old with carcinomas *in situ* ranging from 2mm to 2cm in diameter were recruited, along with 98 healthy controls. They analysed 1 μL of serum for each patient, after diluting it with 3 μL of distilled water and drying it on a Si-plate. Subsequently, the spectra collected were separated into classes using two independent machine learning classification methods: cluster analysis (unsupervised) and artificial neural networks (ANN; supervised). Both models performed with sensitivity and specificity above 95%. Sensitivity and specificity of cluster analysis accounted for 98% and 95% respectively, whilst a sensitivity of 92% and a specificity of 100% were determined with ANN. The results were also tested against 11 other different diseases separately to confirm no other factors were interfering in the classification models and 79% of breast cancer patients were successfully assigned to the correct class.

Ollesch *et al.* [46] investigated the use of HT-FTIR transmission spectroscopy to distinguish between urinary bladder cancer and suspected bladder cancer patients suffering from inflammation. The aim of their research was to identify the spectral biomarkers responsible for the classification and validate them through the study. 135 were the patients included in the dataset: 89 urinary bladder cancer cases and 46 controls, all confirmed by cytology and pathological biopsy. Both linear discriminant analysis (LDA) and random forest (RF) were used as classification models; the latter

performed better, scoring high in sensitivity ($93\pm 10\%$) but poorly in specificity ($46\pm 18\%$); results being attributable to the class imbalance. In 2014, the research team updated their results by presenting multiple data analysis outputs on different datasets, the wider one of 286 patients in total (urinary bladder cancer versus control patients). RF classification outperformed the previous specificity values obtained, accounting for $86\pm 9\%$, however the sensitivity dropped of more than ten percentage points to $82\pm 9\%$ [47].

Maitra *et al.* [48] studied the potential discrimination between four different classes of oesophageal diseases: inflammation, Barrett's, low-grade dysplasia, high-grade dysplasia and oesophageal adenocarcinoma. They analysed plasma, saliva, urine and serum; the genetic algorithms combined with quadratic discriminant analysis (GA-QDA) model achieved a remarkable performance with the serum dataset, scoring only four classification errors in the training set and two in the validation set on a total of 124 samples; this showed a powerful use of the ATR-FTIR technique for oesophageal cancer discrimination. Moreover, Yap *et al.* [49] investigated extracellular vesicles contained in urine as potential biomarker for prostate cancer detection. Preliminary results of a principal component analysis combined with linear discriminant analysis (PCA-LDA) model performed on a small dataset (11 patients) showed a potential development of the ATR-FTIR spectroscopy diagnostic approach for prostate cancer, as the model achieved a sensitivity of 83% and a specificity of 60%. In contrast, Giamougiannis *et al.* [50] analysed ascetic fluids (i.e. abnormal secretion of fluid in the peritoneal cavity) of 45 patients with benign gynaecological pathologies, borderline ovarian tumours and confirmed ovarian cancer. The team performed partial least squares discriminant analysis (PLS-DA) on the spectral dataset, obtaining 79% sensitivity and 93% specificity of the model discriminating cancer samples against benign pathologies and the borderline category; a different data analysis approach compared to the ovarian cancer study published in 2013 by Gajjar *et al.* [42], where the best classification rate ($95\pm 8\%$; i.e. average of sensitivity and specificity percentages) to individuate ovarian cancer by serum ATR-FTIR analysis was achieved by combining forward feature selection (FFS) and the k-nearest neighbours (k-NN) classifier.

Notwithstanding the potential of the analysis of serum with FTIR spectroscopy for the clinical environment, none of these studies have progressed beyond the proof-of-concept stage. However, the aforementioned prospective clinical validation study from Butler *et al.* [31] was recently presented with promising results for the detection of brain tumour. Out of 104 patients referred from their GP for medical imaging, 12 patients had brain tumours: four glioblastoma multiforme (GBM), three anaplastic astrocytoma, two oligoastrocytoma and a medullo-blastoma, an ependymoma and a gliosarcoma. The ATR-FTIR spectra acquired from the prospective group of patients were used to predict the diagnosis. The support vector machine (SVM) classification model was trained with a retrospective cohort of 724 samples. Sensitivity and specificity of the predicted diagnosis of blind data accounted for 83.3% and 87.0% respectively, whilst 93.2% and 92.8% were the outputs related to the retrospective cohort. This study represents the step towards the translation into the clinical environment, having started from a proof-of-principle study in 2014 based on 97 patients [51] and subsequently upgraded to a cohort of 433 patients in 2016, achieving 92.8% of sensitivity and 91.5% of specificity in an optimised RF classification model [40]. An integration of the ATR-FTIR serum analysis into the clinic proposed by Butler *et al.* [31] is showed in Figure 3.

Given the difficulties of diagnosis through imaging, brain tissue biopsies remain the only definitive way to diagnose brain cancer. The risks associated with a brain tumour biopsy include stroke and death, and although the overall risk is small (less than 2%) [52], a non-invasive alternative is highly desirable. Further to the early detection of brain tumours at the primary care stage, Cameron *et al.* [53] have recently evidenced the potential to differentiate between brain tumour types. In this study, 41 lymphoma and 71 GBM serum samples were analysed with ATR-FTIR spectroscopy. The resultant

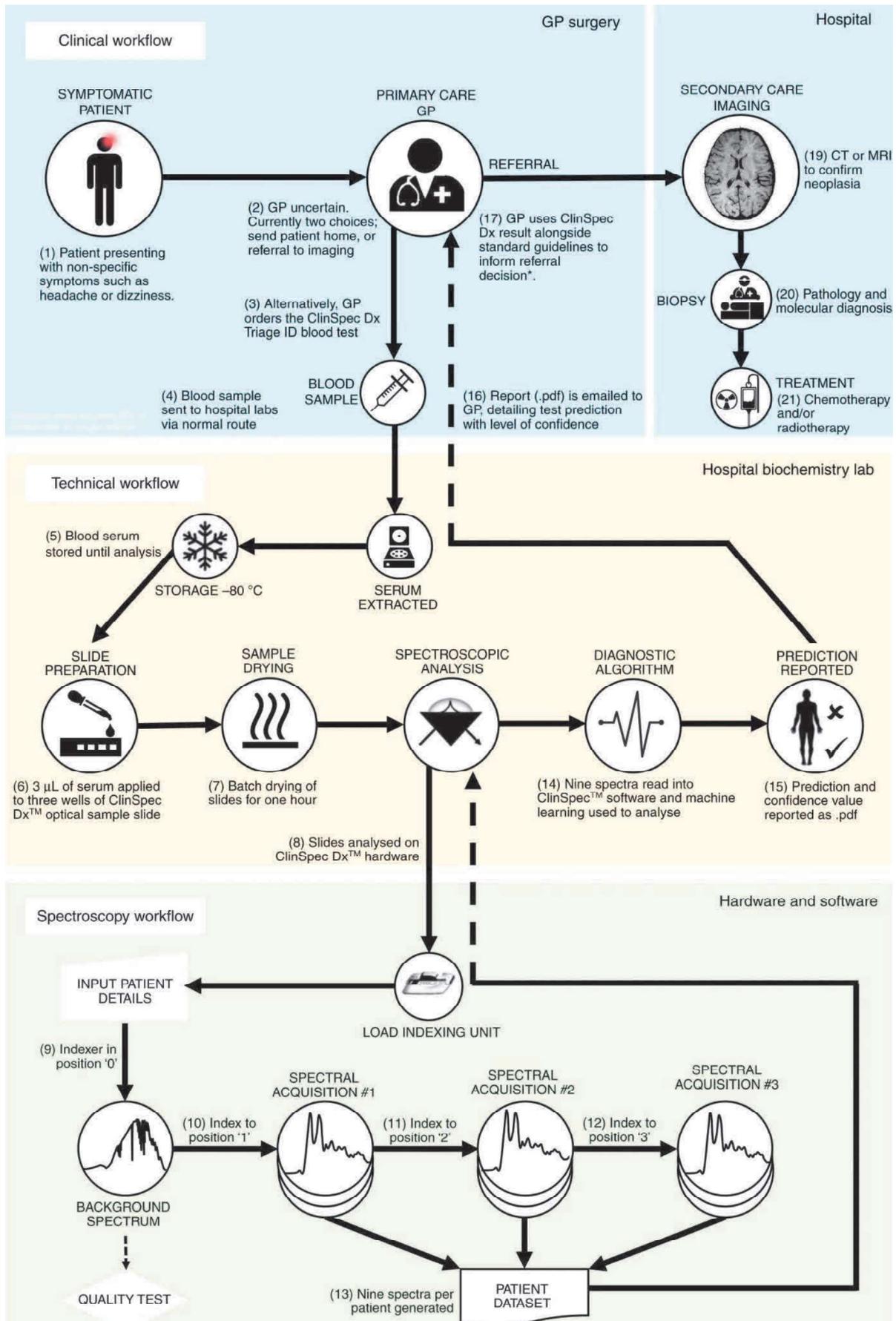


Figure 3 – Proposed integration of a blood test for the triage of brain cancer. Reproduced from ref. [31] (CC by 4.0)

spectra were then analysed with three classification models: RF, PLS-DA and SVM; PLS-DA gave the best outcomes in terms of specificity and sensitivity, at 90.1% and 86.3% respectively.

It is clear that the utility of FTIR spectroscopy in the clinical environment, implemented with machine learning algorithms, could be an extremely valuable option in cancer research to obtain an early diagnosis and discover novel tumour biomarkers. A triage tool consisting in human blood serum analysis with ATR-FTIR spectroscopy, set in primary care, would help GPs in their 'decision to refer' process by providing a comprehensive biochemical panel indicating the likely presence of tumours. ATR-FTIR spectroscopy of serum is a promising, rapid, simple, non-invasive, sensitive, specific and cost-effective cancer diagnosis system [15,31,43].

3. Pre-clinical Considerations

One of the significant benefits of ATR-FTIR spectroscopy as a diagnostic platform is that the technique is reagent free, and therefore does not require extensive sample preparation steps that may otherwise introduce variation into the data. In overview, the sample preparation for serum spectroscopy can be considered as follows: (i) sample collection, (ii) sample storage, (iii) sample deposition, and finally, (iv) sample drying. At each of these steps, however, there is still the possibility of introducing unwanted variance, and understanding these variances is key to developing a consistent and stable diagnostic tool. These pre-clinical considerations have been addressed in the developing literature in the field [54].

Blood collection is a routine process that is largely standardised in clinical practice and research; however, small deviations to collection protocols can have an impact on each sample, and therefore the quality of spectral data [33]. For blood plasma collection, a number of anticoagulants – including heparin and EDTA – are used in clinical practise in order to prevent blood clotting. It has been shown that the choice of anticoagulant is noticeable in the resultant infrared spectrum, which may conceal underlying biological information [55,56]. Whilst the use of serum may overcome this collection issue, other factors (e.g. diet) can also influence the spectral response; however, it is possible to quantitatively monitor the relative levels of dietary components such as glucose, lipids, and even alcohol, using ATR-FTIR spectroscopy and thus, general variance of these factors in the population may need to be considered in cancer diagnostics [57–59].

Sample storage is critical with regards to spectral quality as well as logistically fitting into the potential cancer pathway. Typically, serum samples are allowed to clot, before being temporarily refrigerated at 4°C or frozen at -80°C for later use and long-term storage. The use of frozen samples is often unavoidable for retrospective clinical studies, but it has been suggested that freezing may be beneficial for uptake into the clinical pathway, as sample analysis is then less time dependent [31]. Lovergne *et al.* presented that the largest spectral change in blood components was apparent between fresh and frozen samples, whereas subsequent freeze-thaw cycles had no further spectral impact [56]. This information would suggest that working with frozen samples can be beneficial to sample processing, as well as sample stability. Further studies by these authors have also illustrated that long term storage of samples does not have significant impact on spectral output [54,60].

Inherent to the analysis of blood products by ATR-FTIR spectroscopy is the IRE. In much of the literature, the sample is often deposited on the top surface of the IRE, allowing the evanescent wave to penetrate through the dried, or wet, sample [10]. Other studies have also employed alternative substrates for samples, in order to overcome potential technical and throughput barriers to diagnostic applications [31,61,62]. For both approaches, the method of sample deposition can be highly variable between operators due to differences in pipetting accuracy that can impact sample volume and spread, as well as positional accuracy [56]. A recent study observed that serum deposition upon novel IREs was wildly variable between trained and untrained users and that this could be observed spectrally [54].

Due to the strong IR absorption of water, serum samples are often analysed in the dried state to avoid masking of the fingerprint region by water signals. Whilst liquid serum analysis is possible, more spectral information is visible in dried serum samples [63,64]. Due to the complex composition of blood serum, drying patterns are prominent within dried serum spots, with the coffee ring effect commonly observed as a confounding factor in transmission and transflection FTIR spectroscopy studies [15, 16]. The large sampling area employed in the ATR modality of FTIR spectroscopy, determined by the IRE surface area, helps to overcome spectral variance that can arise from these patterns due macro-measurements of the entire serum spot; however, controlled drying conditions have been shown to improve spectral reproducibility [54,60,65]. The drying environment is crucial to serum drop homogeneity, with humidity, temperature and air flow having significant impact on macromolecule spread in the serum drop, as well as crack features, which are known to cause spectra distortions in all modalities of FTIR spectroscopy [65]. It is clear that sample deposition and drying are therefore closely linked; for instance, serum volume can impact the drying profile of biofluids.

Figure 4 schematises the analytical requisites for clinical validation and introduces the steps required for development of clinical trials, further explained in the following paragraph.

4. Clinical Translation: Requirements and Developments of Clinical Trials

The ultimate goal for laboratory-based research is the translation to the clinic; however, surprisingly few technologies make this transition [66]. The field of biomedical spectroscopy, including ATR-FTIR spectroscopy, has developed considerably in the last two decades yet viable clinical products are only just beginning to emerge [67]. Key barriers to translation include disruption to clinical workflows and increased economic burden on healthcare providers; despite the clinical benefits that new technologies may offer [15].

Generally speaking, defining the clinical question is the first step in translational research and should tackle an unmet need in the treatment, diagnosis or another specific aspect of a disease or clinical problem. The unmet need should be fully explored; for example, in diagnostic applications, the diagnostic pathway should be carefully mapped in order to deduce the underlying issues that need to be resolved [68]. For the early detection of brain cancer, this process was overviewed in a health economic assessment that characterised the current patient pathway and furthermore reported the cost implications of a more efficient pathway [22]. A key aspect of this clinical research is the close interaction with clinicians, key opinion leaders and stakeholders, whose involvement is paramount to defining the clinical utility of any new technologies [15].

Coupling this clinical question to an appropriate technique or technology is the next consideration. This may require technical developments or adaptations in order to best suit the clinical question, with automation a desirable feature of any new technology [69]. For translation ATR-FTIR spectroscopy, novel point-of-care technologies including handheld and microfluidic devices have been discussed, as well as alternative IRE configurations for batch processing [28,31,70].

Proof-of-concept studies are essential evidence of the utility of any new approach to a clinical question and will begin to define the potential benefit that can be brought to the clinical pathway. As previously discussed, a wealth of proof-of-concept studies have been already completed. For cancer diagnostics, often the first step is a binary classification of disease *versus* non-diseased patients, with samples usually obtained from retrospectively stored biobanks. Sample numbers are often a limiting factor in the development of new technologies, and a noticeable pattern in the literature is an alarming low number of patient numbers used to establish early utility. The use of machine learning approaches often employed to classify between patients is heavily influenced by the number of patients fed into the training and testing of these computational models [71]. Whilst these studies are useful at identifying potential use, often performance indicators are not reflective of true clinical performance. It has been suggested that sample sizes in the range of 75-100 samples per class may

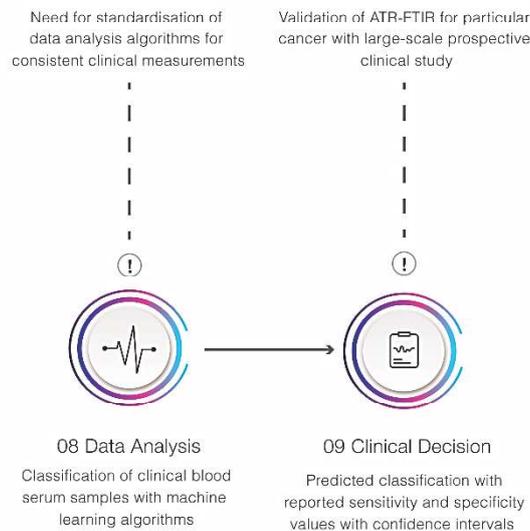
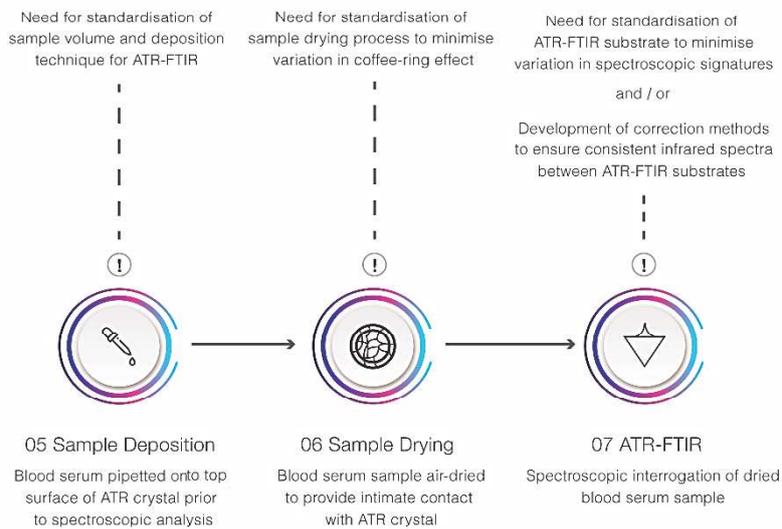
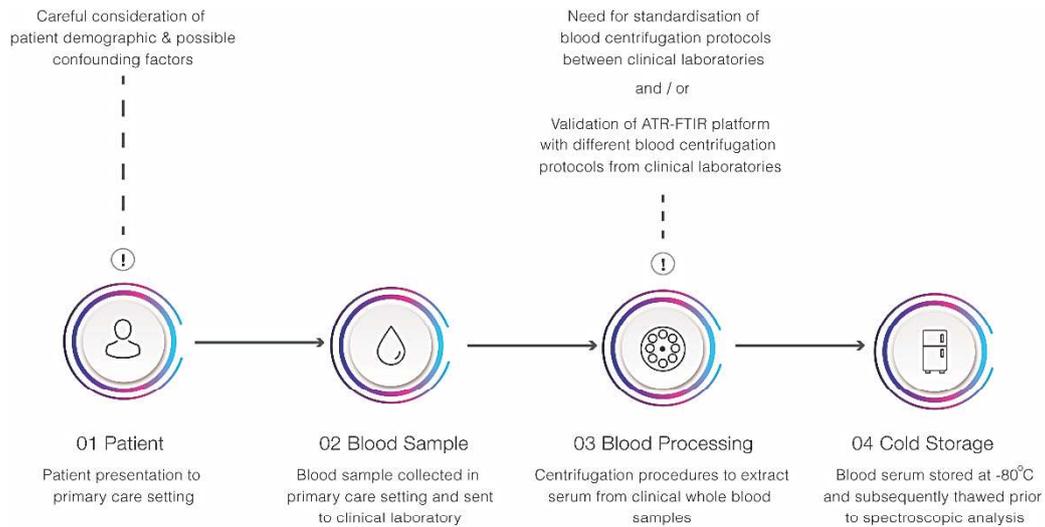


Figure 4 - Stepwise process of the current pre-analytical path, followed by analysis steps and clinical decision.

be sufficient to establish good, but not perfect, classifiers at the proof-of-concept stage [72]. More complex studies can be defined by observing other samples classes, such as expanding from binary cancer *versus* non-cancer questions, to further stratification with other diseases and disease sub-types, such as prostate cancer grading [27].

Determining the true clinical utility may only be possible through studies conducted in the targeted clinical environment and patient cohort. Clinical feasibility studies, or diagnostic accuracy studies, may be required in order to ascertain the diagnostic performance of new tests and technologies. The performance can be measured depending on the type of study, with cancer diagnostics often measured in terms of sensitivity and specificity; the ability to either accurately detect patients with disease or those without, respectively. Although, it is important to note that other performance characteristics may be important to clinical utility based on specific questions (e.g. population screening), as disease prevalence can have a significant impact on positive and negative predictive values, as well as sample numbers required for powered clinical studies. Such trials are required for any downstream regulatory process, in which new technologies will require clear trial design and documented validation throughout the process. A crucial point in spectroscopic cancer diagnostics is that any tests or devices that require regulatory approval will need to be standardised and used throughout any pivotal clinical studies, including any machine learning approaches. Prospective clinical trials can be costly, but also provide the highest level of evidence.

In addition to scientific and clinical development, there is also the commercial element; intellectual property (IP), funding and international markets can often represent a further barrier to translation.

5. Conclusions

Early detection of cancer is vital for improving of survival rates and quality of life. A simple, rapid, cost-effective, sensitive and specific test is needed in the clinical environment to achieve early detection, avoid overtreating and over-diagnosing patients, and to reduce morbidity and mortality. ATR-FTIR analysis of biofluids, particularly serum, has been seen to represent an important tool in cancer research thanks to the impressive statistical results obtained with different machine learning algorithms. However, despite the numerous promising proof-of-concept studies performed in the last decades, clinical translation has yet to be achieved. A prospective clinical validation study on brain tumours has been recently published [31], carrying hopes and expectations for the clinical spectroscopy community.

Notwithstanding the continuous improvements in cancer research, thanks to the developments of instrumentation and analytical methods, the success of clinical translation cannot be confined to the scientific world; it also need to be considered under the commercial view. Intellectual property matters, difficulties in obtaining funding and competition on international markets often represent barriers for proof-of-concept studies, making the ultimate goal difficult to be achieved. The future development of clinical spectroscopy towards clinical translation is growing rapidly and poised for a tipping point. The development of clinically relevant methodologies, hardware and studies combined with commercial and clinical pull is enabling the technique to find a rapid translational path. Within the next five years we would expect a clinical spectroscopic technique in use for patient benefit.

Corresponding Author

*E-mail: matthew.baker@strath.ac.uk – matthew.baker@clinspecdx.com

*Twitter: @ChemistryBaker – @clinspecdx

Declaration of Competing Interest

MJB and HJB are directors of ClinSpec Diagnostics Ltd and PMB is a consultant for ClinSpec Diagnostics Ltd.

Acknowledgements

The authors would like to thank the EPSRC (EP/L505080/1), Scottish Enterprise and ClinSpec Diagnostics Ltd for funding.

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA. Cancer J. Clin.* 68 (2018) 394–424. <https://doi.org/10.3322/caac.21492>.
- [2] GLOBOCAN 2018, Cancer Statistics: Cancer Tomorrow (2018-2040). <https://gco.iarc.fr/tomorrow/home> (accessed December 5, 2019).
- [3] WHO, The top 10 causes of death, (2018). <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death> (accessed December 4, 2019).
- [4] C. Sawyers, Targeted cancer therapy, *Nature.* 432 (2004) 294–297. <https://doi.org/10.1038/nature03095>.
- [5] J.D. Schiffman, P.G. Fisher, P. Gibbs, Early Detection of Cancer: Past, Present, and Future, *Am. Soc. Clin. Oncol. Educ. B.* 35 (2015) 57–65. https://doi.org/10.14694/EdBook_AM.2015.35.57.
- [6] G. Sölétormos, M.J. Duffy, S. Othman Abu Hassan, R.H.M. Verheijen, B. Tholander, R.C. Bast, K.N. Gaarenstroom, C.M. Sturgeon, J.M. Bonfrer, P.H. Petersen, H. Troonen, G. Carlotorre, J. Kanty Kulpa, M.K. Tuxen, R. Molina, Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines from the European Group on Tumor Markers, in: *Int. J. Gynecol. Cancer*, Lippincott Williams and Wilkins, 2016: pp. 43–51. <https://doi.org/10.1097/IGC.0000000000000586>.
- [7] K. Mistry, G. Cable, Meta-analysis of prostate-specific antigen and digital rectal examination as screening tests for prostate carcinoma, *J. Am. Board Fam. Pract.* 16 (2003) 95–101. <https://doi.org/10.3122/jabfm.16.2.95>.
- [8] U.K. Ballehaninna, R.S. Chamberlain, The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal, *J. Gastrointest. Oncol.* 3 (2012) 105–119. <https://doi.org/10.3978/j.issn.2078-6891.2011.021>.
- [9] C. Fiala, E.P. Diamandis, Utility of circulating tumor DNA in cancer diagnostics with emphasis on early detection, *BMC Med.* 16 (2018) 166. <https://doi.org/10.1186/s12916-018-1157-9>.
- [10] L. Lechowicz, M. Chrapek, J. Gaweda, M. Urbaniak, I. Konieczna, Use of Fourier-transform infrared spectroscopy in the diagnosis of rheumatoid arthritis: a pilot study, *Mol. Biol. Rep.* 43 (2016) 1321–1326. <https://doi.org/10.1007/s11033-016-4079-7>.
- [11] A. Travo, C. Paya, G. Déléris, J. Colin, B. Mortemousque, I. Forfar, Potential of FTIR spectroscopy for analysis of tears for diagnosis purposes, *Anal. Bioanal. Chem.* 406 (2014) 2367–2376. <https://doi.org/10.1007/s00216-013-7607-5>.
- [12] J. Titus, E. Viennois, D. Merlin, A.G. Unil Perera, Minimally invasive screening for colitis using attenuated total internal reflectance fourier transform infrared spectroscopy, *J. Biophotonics.* 10 (2017) 465–472. <https://doi.org/10.1002/jbio.201600041>.
- [13] V. Untereiner, G. Dhruvananda Sockalingum, R. Garnotel, C. Gobinet, F. Ramaholimihaso, F. Ehrhard, M.D. Diebold, G. Thiéfin, Bile analysis using high-throughput FTIR spectroscopy for the diagnosis of malignant biliary strictures: A pilot study in 57 patients, *J. Biophotonics.* 7 (2014) 241–253. <https://doi.org/10.1002/jbio.201300166>.

- [14] S. Roy, D. Perez-Guaita, D.W. Andrew, J.S. Richards, D. McNaughton, P. Heraud, B.R. Wood, Simultaneous ATR-FTIR Based Determination of Malaria Parasitemia, Glucose and Urea in Whole Blood Dried onto a Glass Slide, *Anal. Chem.* 89 (2017) 5238–5245. <https://doi.org/10.1021/acs.analchem.6b04578>.
- [15] D. Finlayson, C. Rinaldi, M.J. Baker, Is Infrared Spectroscopy Ready for the Clinic?, *Anal. Chem.* 91 (2019) 12117–12128. <https://doi.org/10.1021/acs.analchem.9b02280>.
- [16] J. Greener, B. Abbasi, E. Kumacheva, Attenuated total reflection Fourier transform infrared spectroscopy for on-chip monitoring of solute concentrations, *Lab Chip.* 10 (2010) 1561. <https://doi.org/10.1039/c001889a>.
- [17] Perkin Elmer, FT-IR Spectroscopy: Attenuated Total Reflectance (ATR), (n.d.). https://web.archive.org/web/20070216065646/http://las.perkinelmer.com/content/TechnicalInfo/TCH_FTIRATR.pdf (accessed November 21, 2019).
- [18] K.M. Dorling, M.J. Baker, Highlighting attenuated total reflection Fourier transform infrared spectroscopy for rapid serum analysis, *Trends Biotechnol.* 31 (2013) 327–328. <https://doi.org/10.1016/j.tibtech.2013.03.010>.
- [19] Bruker Optics, ATR: Advantages for FT-IR Spectroscopy, (n.d.). <https://www.azom.com/article.aspx?ArticleID=5958> (accessed November 21, 2019).
- [20] A.C.S. Talari, M.A.G. Martinez, Z. Movasaghi, S. Rehman, I.U. Rehman, Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues, *Appl. Spectrosc. Rev.* 52 (2017) 456–506. <https://doi.org/10.1080/05704928.2016.1230863>.
- [21] Z. Movasaghi, S. Rehman, D.I. ur Rehman, Fourier Transform Infrared (FTIR) Spectroscopy of Biological Tissues, *Appl. Spectrosc. Rev.* 43 (2008) 134–179. <https://doi.org/10.1080/05704920701829043>.
- [22] E. Gray, H.J. Butler, R. Board, P.M. Brennan, A.J. Chalmers, T. Dawson, J. Goodden, W. Hamilton, M.G. Hegarty, A. James, M.D. Jenkinson, D. Kernick, E. Lekka, L.J. Livermore, S.J. Mills, K. O'Neill, D.S. Palmer, B. Vaqas, M.J. Baker, Health economic evaluation of a serum-based blood test for brain tumour diagnosis: exploration of two clinical scenarios, *BMJ Open.* 8 (2018) e017593. <https://doi.org/10.1136/bmjopen-2017-017593>.
- [23] M.J. Walsh, S.E. Holton, A. Kajdacsy-Balla, R. Bhargava, Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology, *Vib. Spectrosc.* 60 (2012) 23–28. <https://doi.org/10.1016/j.vibspec.2012.01.010>.
- [24] B. Bird, M. Miljković, S. Remiszewski, A. Akalin, M. Kon, M. Diem, Infrared spectral histopathology (SHP): a novel diagnostic tool for the accurate classification of lung cancer, *Lab. Investig.* 92 (2012) 1358–1373. <https://doi.org/10.1038/labinvest.2012.101>.
- [25] P. Lasch, W. Haensch, D. Naumann, M. Diem, Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis, *Biochim. Biophys. Acta - Mol. Basis Dis.* 1688 (2004) 176–186. <https://doi.org/10.1016/j.bbadis.2003.12.006>.
- [26] M.J. Baker, E. Gazi, M.D. Brown, J.H. Shanks, P. Gardner, N.W. Clarke, FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer, *Br. J. Cancer.* 99 (2008) 1859–1866. <https://doi.org/10.1038/sj.bjc.6604753>.
- [27] E. Gazi, M. Baker, J. Dwyer, N.P. Lockyer, P. Gardner, J.H. Shanks, R.S. Reeve, C.A. Hart, N.W. Clarke, M.D. Brown, A Correlation of FTIR Spectra Derived from Prostate Cancer Biopsies with Gleason Grade and Tumour Stage, *Eur. Urol.* 50 (2006) 750–761. <https://doi.org/10.1016/j.eururo.2006.03.031>.

- [28] 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, *J. Biophotonics*. 7 (2014) 153–165. <https://doi.org/10.1002/jbio.201400018>.
- [29] M.J. Baker, S.R. Hussain, L. Lovergne, V. Untereiner, C. Hughes, R.A. Lukaszewski, G. Thiéfin, G.D. Sockalingum, Developing and understanding biofluid vibrational spectroscopy: A critical review, *Chem. Soc. Rev.* 45 (2016) 1803–1818. <https://doi.org/10.1039/c5cs00585j>.
- [30] A.R. Shaw, S. Low-Ying, A. Man, K.Z. Liu, C. Mansfield, C.B. Rileg, M. Vijarnsorn, Infrared Spectroscopy of Biofluids in Clinical Chemistry and Medical Diagnostics, in: *Biomed. Vib. Spectrosc.*, John Wiley & Sons, Inc., 2007: pp. 79–103. <https://doi.org/10.1002/9780470283172.ch4>.
- [31] H.J. Butler, P.M. Brennan, J.M. Cameron, D. Finlayson, M.G. Hegarty, M.D. Jenkinson, D.S. Palmer, B.R. Smith, M.J. Baker, Development of high-throughput ATR-FTIR technology for rapid triage of brain cancer, *Nat. Commun.* 10 (2019) 1–9. <https://doi.org/10.1038/s41467-019-12527-5>.
- [32] H.J. Butler, B.R. Smith, R. Fritzsche, P. Radhakrishnan, D.S. Palmer, M.J. Baker, Optimised spectral pre-processing for discrimination of biofluids via ATR-FTIR spectroscopy, *Analyst*. 143 (2018) 6121–6134. <https://doi.org/10.1039/C8AN01384E>.
- [33] M.K. Tuck, D.W. Chan, D. Chia, A.K. Godwin, W.E. Grizzle, K.E. Krueger, W. Rom, M. Sanda, L. Sorbara, S. Stass, W. Wang, D.E. Brenner, Standard Operating Procedures for Serum and Plasma Collection: Early Detection Research Network Consensus Statement *Standard Operating Procedure Integration Working Group*, *J. Proteome Res.* 8 (2009) 113–117. <https://doi.org/10.1021/pr800545q>.
- [34] F. Bonnier, M.J. Baker, H.J. Byrne, Vibrational spectroscopic analysis of body fluids: avoiding molecular contamination using centrifugal filtration, *Anal. Methods*. 6 (2014) 5155. <https://doi.org/10.1039/c4ay00891j>.
- [35] R.S. Tirumalai, K.C. Chan, D.R.A. Prieto, H.J. Issaq, T.P. Conrads, T.D. Veenstra, Characterization of the low molecular weight human serum proteome., *Mol. Cell. Proteomics*. 2 (2003) 1096–1103. <https://doi.org/10.1074/mcp.M300031-MCP200>.
- [36] E.F. Petricoin, C. Belluco, R.P. Araujo, L.A. Liotta, The blood peptidome: A higher dimension of information content for cancer biomarker discovery, *Nat. Rev. Cancer*. 6 (2006) 961–967. <https://doi.org/10.1038/nrc2011>.
- [37] I. Infusino, M. Panteghini, Serum albumin: Accuracy and clinical use, *Clin. Chim. Acta.* 419 (2013) 15–18. <https://doi.org/10.1016/j.cca.2013.01.005>.
- [38] F. Bonnier, G. Brachet, R. Duong, T. Sojinrin, R. Respaud, N. Aubrey, M.J. Baker, H.J. Byrne, I. Chourpa, Screening the low molecular weight fraction of human serum using ATR-IR spectroscopy, *J. Biophotonics*. 9 (2016) 1085–1097. <https://doi.org/10.1002/jbio.201600015>.
- [39] F. Bonnier, H. Blasco, C. Wasselet, G. Brachet, R. Respaud, L.F.C.S. Carvalho, D. Bertrand, M.J. Baker, H.J. Byrne, I. Chourpa, Ultra-filtration of human serum for improved quantitative analysis of low molecular weight biomarkers using ATR-IR spectroscopy, *Analyst*. 142 (2017) 1285–1298. <https://doi.org/10.1039/c6an01888b>.
- [40] 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 differentiation: rapid stratified serum diagnostics via attenuated total reflection Fourier-transform infrared spectroscopy, *J.*

- Neurooncol. 127 (2016) 463–472. <https://doi.org/10.1007/s11060-016-2060-x>.
- [41] X. Sun, Y. Xu, J. Wu, Y. Zhang, K. Sun, Detection of lung cancer tissue by attenuated total reflection-Fourier transform infrared spectroscopy - A pilot study of 60 samples, *J. Surg. Res.* 179 (2013) 33–38. <https://doi.org/10.1016/j.jss.2012.08.057>.
- [42] 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. <https://doi.org/10.1039/c3an36654e>.
- [43] A.A. Bunaciu, V.D. Hoang, H.Y. Aboul-Enein, Applications of FT-IR Spectrophotometry in Cancer Diagnostics, *Crit. Rev. Anal. Chem.* 45 (2015) 156–165. <https://doi.org/10.1080/10408347.2014.904733>.
- [44] A.A. Bunaciu, Ş. Fleschin, V.D. Hoang, H.Y. Aboul-Enein, Vibrational Spectroscopy in Body Fluids Analysis, *Crit. Rev. Anal. Chem.* 47 (2017) 67–75. <https://doi.org/10.1080/10408347.2016.1209104>.
- [45] J. Backhaus, R. Mueller, N. Formanski, N. Szlama, H.-G. Meerpohl, M. Eidt, P. Bugert, Diagnosis of breast cancer with infrared spectroscopy from serum samples, *Vib. Spectrosc.* 52 (2010) 173–177. <https://doi.org/10.1016/j.vibspec.2010.01.013>.
- [46] J. Ollesch, S.L. Drees, H.M. Heise, T. Behrens, T. Brüning, K. Gerwert, FTIR spectroscopy of biofluids revisited: an automated approach to spectral biomarker identification, *Analyst.* 138 (2013) 4092. <https://doi.org/10.1039/c3an00337j>.
- [47] J. Ollesch, M. Heinze, H.M. Heise, T. Behrens, T. Brüning, K. Gerwert, It's in your blood: spectral biomarker candidates for urinary bladder cancer from automated FTIR spectroscopy, *J. Biophotonics.* 7 (2014) 210–221. <https://doi.org/10.1002/jbio.201300163>.
- [48] I. Maitra, C.L.M. Morais, K.M.G. Lima, K.M. Ashton, R.S. Date, F.L. Martin, Attenuated total reflection Fourier-transform infrared spectral discrimination in human bodily fluids of oesophageal transformation to adenocarcinoma, *Analyst.* (2019). <https://doi.org/10.1039/C9AN01749F>.
- [49] X. L. Yap, T. A. Ong, J. Lim, B. Wood, W. L. Lee, A pilot study of prostate cancer-derived extracellular vesicles in urine using IR spectroscopy, *Prog. Drug Discov. Biomed. Sci.* 2 (2019) 16–19. <http://www.journals.hh-publisher.com/index.php/pddbs/article/view/56>.
- [50] P. Giamougiannis, R. Grabowska, C. Morais, A. Anagnostopoulos, L. Whitham, E. Tingi, G. Owens, P. Hadjiyiannakis, N. Wood, P. Martin-Hirsch, F. Martin, EP857 Distinguishing benign vs. cancer states in ovary based on spectrochemical analysis of ascites: a budget omics approach, *Int. J. Gynecol. Cancer.* 29 (2019) A466–A467. <https://doi.org/10.1136/ijgc-2019-esgo.906>.
- [51] 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, *J. Biophotonics.* 7 (2014) 189–199. <https://doi.org/10.1002/jbio.201300149>.
- [52] R.L. Yong, R.R. Lonser, Safety of closed brain biopsy: Population-based studies weigh in, *World Neurosurg.* 79 (2013) 53–54. <https://doi.org/10.1016/j.wneu.2012.05.016>.
- [53] J.M. Cameron, H.J. Butler, B.R. Smith, M.G. Hegarty, M.D. Jenkinson, K. Syed, P.M. Brennan, K. Ashton, T. Dawson, D.S. Palmer, M.J. Baker, Developing infrared spectroscopic detection for stratifying brain tumour patients: Glioblastoma multiforme: Vs. lymphoma, *Analyst.* 144

- (2019) 6736–6750. <https://doi.org/10.1039/c9an01731c>.
- [54] J.M. Cameron, Manuscript in Preparation.
- [55] M. Martin, D. Perez-Guaita, D.W. Andrew, J.S. Richards, B.R. Wood, P. Heraud, The effect of common anticoagulants in detection and quantification of malaria parasitemia in human red blood cells by ATR-FTIR spectroscopy, *Analyst*. 142 (2017) 1192–1199. <https://doi.org/10.1039/C6AN02075E>.
- [56] L. Lovergne, P. Bouzy, V. Untereiner, R. Garnotel, M.J. Baker, G. Thiéfin, G.D. Sockalingum, Biofluid infrared spectro-diagnostics: Pre-analytical considerations for clinical applications, *Faraday Discuss.* 187 (2016) 521–537. <https://doi.org/10.1039/c5fd00184f>.
- [57] D. Perez-Guaita, J. Ventura-Gayete, C. Pérez-Rambla, M. Sancho-Andreu, S. Garrigues, M. de la Guardia, Evaluation of infrared spectroscopy as a screening tool for serum analysis, *Microchem. J.* 106 (2013) 202–211. <https://doi.org/10.1016/j.microc.2012.06.016>.
- [58] A. Oleszko, J. Hartwich, A. Wójtowicz, M. Gąsior-Głogowska, H. Huras, M. Komorowska, Comparison of FTIR-ATR and Raman spectroscopy in determination of VLDL triglycerides in blood serum with PLS regression, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 183 (2017) 239–246. <https://doi.org/10.1016/j.saa.2017.04.020>.
- [59] K. Sharma, S.P. Sharma, S.C. Lahiri, Estimation of blood alcohol concentration by horizontal attenuated total reflectance-Fourier transform infrared spectroscopy, *Alcohol*. 44 (2010) 351–357. <https://doi.org/10.1016/j.alcohol.2009.11.004>.
- [60] L. Lovergne, J. Lovergne, P. Bouzy, V. Untereiner, M. Offroy, R. Garnotel, G. Thiéfin, M.J. Baker, G.D. Sockalingum, Investigating pre-analytical requirements for serum and plasma based infrared spectro-diagnostic, *J. Biophotonics*. (2019). <https://doi.org/10.1002/jbio.201900177>.
- [61] 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, *Biotechnol. Prog.* 31 (2015) 832–839. <https://doi.org/10.1002/btpr.2084>.
- [62] S. Roy, D. Perez-Guaita, D.W. Andrew, J.S. Richards, D. McNaughton, P. Heraud, B.R. Wood, Simultaneous ATR-FTIR Based Determination of Malaria Parasitemia, Glucose and Urea in Whole Blood Dried onto a Glass Slide, *Anal. Chem.* 89 (2017) 5238–5245. <https://doi.org/10.1021/acs.analchem.6b04578>.
- [63] A. Sala, Manuscript in Preparation.
- [64] K.M. Elkins, Rapid Presumptive “Fingerprinting” of Body Fluids and Materials by ATR FT-IR Spectroscopy, *J. Forensic Sci.* 56 (2011) 1580–1587. <https://doi.org/10.1111/j.1556-4029.2011.01870.x>.
- [65] J.M. Cameron, H.J. Butler, D.S. Palmer, M.J. Baker, Biofluid spectroscopic disease diagnostics: A review on the processes and spectral impact of drying, *J. Biophotonics*. 11 (2018) 1–12. <https://doi.org/10.1002/jbio.201700299>.
- [66] H.J. Byrne, M. Baranska, G.J. Puppels, N. Stone, B. Wood, K.M. Gough, P. Lasch, P. Heraud, J. Sulé-Suso, G.D. Sockalingum, Spectro-pathology for the next generation: Quo vadis?, *Analyst*. 140 (2015) 2066–2073. <https://doi.org/10.1039/C4AN02036G>.
- [67] H.J. Butler, J.M. Cameron, C.A. Jenkins, G. Hithell, S. Hume, N.T. Hunt and M.J. Baker, Shining a light on clinical spectroscopy: Translation of diagnostic IR, 2D-IR and Raman spectroscopy

towards the clinic, *Clinical Spectroscopy*, in press (2019).

- [68] E. Badrick, K. Cresswell, P. Ellis, A.G. Renehan, E.J. Crosbie, P. Crosbie, P.S. Hall, H. O'Flynn, R. Martin, J. Leighton, L. Brown, D. Makin, R.L. Morris, E. Thorpe, J.C. Dickinson, G. Buckley, Top ten research priorities for detecting cancer early, *Lancet Public Heal.* 4 (2019) e551. [https://doi.org/10.1016/S2468-2667\(19\)30185-9](https://doi.org/10.1016/S2468-2667(19)30185-9).
- [69] C. Kendall, M. Isabelle, F. Bazant-Hegemark, J. Hutchings, L. Orr, J. Babrah, R. Baker, N. Stone, Vibrational spectroscopy: A clinical tool for cancer diagnostics, *Analyst.* 134 (2009) 1029–1045. <https://doi.org/10.1039/b822130h>.
- [70] D. Perez-Guaita, K.M. Marzec, A. Hudson, C. Evans, T. Chernenko, C. Matthäus, M. Miljkovic, M. Diem, P. Heraud, J.S. Richards, D. Andrew, D.A. Anderson, C. Doerig, J. Garcia-Bustos, D. McNaughton, B.R. Wood, Parasites under the Spotlight: Applications of Vibrational Spectroscopy to Malaria Research, *Chem. Rev.* 118 (2018) 5330–5358. <https://doi.org/10.1021/acs.chemrev.7b00661>.
- [71] H.M. Kalayeh, D.A. Landgrebe, Predicting the Required Number of Training Samples, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-5 (1983) 664–667. <https://doi.org/10.1109/TPAMI.1983.4767459>.
- [72] C. Beleites, U. Neugebauer, T. Bocklitz, C. Krafft, J. Popp, Sample size planning for classification models, *Anal. Chim. Acta.* 760 (2013) 25–33. <https://doi.org/10.1016/j.aca.2012.11.007>.