Clinical Spectroscopy 4 (2022) 100020

Contents lists available at ScienceDirect

Clinical Spectroscopy

journal homepage: www.sciencedirect.com/journal/clinical-spectroscopy

An observational cohort study to evaluate the use of serum Raman spectroscopy in a rapid diagnosis center setting

Freya E.R. Woods [a,*], Susan Chandler [b], Natalia Sikora [a], Rachel Harford [c], Ahmad Souriti [c], Helen Gray [c], Heather Wilkes [c], Catherine Lloyd-Bennett [c], Dean A. Harris [b,c], Peter R. Dunstan [a]

[a] Physics Department, College of Science, Centre for NanoHealth, Swansea University, SA2 8PP, United Kingdom
[b] Swansea University Medical School, Swansea University, SA2 8PP, United Kingdom
[c] Swansea Bay University Health Board, Port Talbot SA12 7BR, United Kingdom

ARTICLE INFO

Keywords:
Raman Spectroscopy
Cancer
Rapid Diagnosis Center, Non-specific symptoms
Diagnostics

ABSTRACT

Cancer presenting with non-specific vague symptoms remains a clinical challenge. The purpose of this study was to assess the feasibility of serum Raman spectroscopy for cancer detection in a rapid diagnosis center (RDC) setting. The primary aim was to identify significant spectral peaks of change in sera from cancer patients and the secondary aim was to assign molecular species at Raman peaks.

In this prospective observation study of a secondary care RDC, patients referred with vague cancer-related symptoms were recruited. Raman spectra of blood sera of 54 patients was obtained. Of these, 10 patients were diagnosed with cancer, and 44 no significant pathology (control). Common spectral increase/decrease between control and cancer was seen in spectral peaks 830 cm^{-1}, 878 cm^{-1}, 1031 cm^{-1}, 1174 cm^{-1}, 1397 cm^{-1} tentatively attributed to amino acids, carbohydrates, fatty acids, and proteins. Individual differences between cancer and control via statistical analysis identifies 3 peaks with significance for all 10 of the cancer patients. The peaks are 878 cm^{-1}, 1449 cm^{-1} and 1519 cm^{-1}, tentatively attributed to proteins, amino acids, lipids, fatty acids, glycoproteins, carbohydrates, and carotenoids. Differences are also seen for at least 9 of the cancers in the peaks at 830 cm^{-1}, 851 cm^{-1}, 1127 cm^{-1}, 1174 cm^{-1}, 1270 cm^{-1}, and 1656 cm^{-1}, tentatively attributed to amino acids, lactate, lipids, triglycerides, carbohydrates, and proteins.

Raman spectroscopy has the potential to enhance RDC referral criteria through the detection of peak differences seen commonly with different cancer types. Development of Artificial Intelligence (AI) based models could enable rapid detection and discrimination of different cancer types with more data availability.

1. Introduction

Diagnosis of cancer at an early stage is known to improve survival. The urgent suspected cancer (USC) pathways were introduced across the UK to standardize referrals and investigations for suspected cancer in an effort to reduce time to diagnosis. The National Institute for Health and Care Excellence (NICE) guidelines for USC pathways are based upon cancer site-specific criteria including age, 'red flag' symptoms or clinical signs. Many patients do not fulfill the NICE criteria due to vague and non-specific symptoms. Studies have shown up to 50% of patients diagnosed with cancer do not present to primary care with pathway-specific red flag symptoms that fulfill the current NICE criteria. These patients wait a median of 34 days longer for a diagnosis compared with patients presenting with symptoms that fulfill the criteria [1]. The ongoing global COVID-19 pandemic has placed a huge burden upon healthcare and led to delays in diagnostic pathways with consequent effects on cancer outcomes [2]. This threatens the goals of the NHS Long Term Plan which seeks to diagnose 75% of cancers at an earlier stage (I/II) by 2028 and to facilitate the faster diagnostic standard of 28 days [3].

Rapid Diagnosis Centers (RDCs) provide an opportunity for detection of cancer when site specific symptoms are absent. Based on a similar clinic model developed in Denmark [4] two RDCs are currently established in Wales with planned roll-out in all health boards. A study of RDC activity in Swansea Bay University Health Board between 2017 and 2018 showed that the mean time to diagnosis was just 5.9 days when diagnosis was made through the clinic compared to 84.2 days in usual care [5]. Above 80% capacity the RDC produced more quality-adjusted

* Corresponding author.
E-mail address: f.woods.795986@swansea.ac.uk (F.E.R. Woods).

https://doi.org/10.1016/j.clispe.2022.100020
Received 16 September 2021; Received in revised form 21 December 2021; Accepted 18 January 2022
Available online 20 January 2022

life years (QALYs) and was less costly than standard clinical practice. High levels of patient and GP satisfaction with the RDC service are reported in Wales [6]. In parallel this model is being rolled out nationally by NHS England and NHS Improvement. Pilot Multi-Disciplinary Centers (MDCs) were established to improve cancer outcomes within the Accelerate Coordinate Evaluate (ACE) program [7] with at least one Rapid Access Center mandated per cancer alliance with full population coverage by 2024 [8].

Due to the nature of referrals with non-specific symptoms there is a low conversion rate to a cancer diagnosis observed within RDCs (7–12%). It would be beneficial to develop simple blood-based triage tests to align with the top ten priorities of early cancer detection and with the RDC model [9].

Raman spectroscopy (RS) is a vibrational spectroscopic technique that provides rapid analysis of biological samples. RS simultaneously measures a range of molecular species (including proteins, nucleic acids, and lipids) within biological samples to produce a spectrum or 'chemical fingerprint' unique to the sample. RS techniques have been developed with the ability to detect disease in both tissue and serum for a range of cancers including but not limited to leukemia, cervical, colorectal, oral squamous cell, and breast cancer [10–14]. The advantage of serum is the ease of collection, storage and transport allowing it to be analysed centrally. The ability of RS to detect many different types of cancer could be utilized in an RDC setting with a serum test potentially allowing a reduction in the number of diagnostic investigations needed or more targeted investigations to be carried out and reduce time to diagnosis.

The aims of this study were to explore the potential of RS to identify significant peaks changes in serum from RDC cancer patients and assign Raman peaks with a view to the future development of cancer diagnosis site models.

2. Methods

2.1. Patient recruitment

Full ethical approval for all aspects of this study was granted by the Wales Research and Ethics Committee (REC reference 14/WA/0028). Patients referred to RDC clinic between February 2018 and March 2020 were recruited prior to their initial consultation and a fasting blood sample was obtained by trained staff. Patient demographic data, diagnosis and outcomes were recorded. Exclusion criteria was patients unwilling or unable to give consent or from vulnerable groups. Clinical diagnosis and patient outcomes were obtained from electronic medical records. Cancer diagnosis was confirmed with histopathology. Control patients were followed up for 6 months to ensure no further significant diagnosis were made and to ensures no cancer diagnosis were missed. RDC researchers were blinded to patient diagnosis and outcomes during data collection. Age, gender, polypharmacy (> 5 medications), smoking status, presence of cardiovascular or hepatic disease, and overall diagnosis for patients is shown in Table 1.

2.2. Rapid diagnosis center

Patients fulfilling referral criteria (unexplained weight loss, pain, fatigue, or shortness of breath) are referred from their GP to the RDC and are seen within a 2-week target. The service is composed of a multidisciplinary team of consultant physician, radiologist, clinical nurse specialist and healthcare support workers. At referral, the GP arranges a panel of routine blood tests, tumor markers and a chest radiograph. Patients are evaluated in the clinic by history and clinical examination then computed tomography (CT) scan of chest abdomen and pelvis. Results are evaluated on the same day by the MDT before the patient is informed of the outcome: (1) cancer diagnosis with referral to specialist, (2) non-cancer diagnosis, (3) no serious pathology found discharge back to GP, and (4) no diagnosis, continue investigations [5].

Table 1
Demographic and clinical data. 4 patients with significant other pathology have been excluded from demographic table and further analysis., * 1 Patient was considered too frail to undergo further investigation following initial CT scan in RDC.

Total	n = 54		
Demographic data	Total	Cancer	Control
Male (%)	35 (65%)	6	29
Female (%)	19 (35%)	4	15
Age median (range)	71 (28–90) years		
Polypharmacy (on ≥5 medications) (%)	36 (67%)		
Smoking status	**Total**	Cancer	Control
Non-smoker (%)	**29 (54%)**	4	25
Ex-smoker (%)	**20 (37%)**	5	15
Current smoker (%)	5 **(9%)**	1	4
Other health issues	**Total**	Cancer	Control
Cardiovascular disease	**37 (69%)**	7 (70%)	30 (68%)
Hepatic disease	**3 (6%)**	0 (0%)	3 (7%)
Diagnosis			
Cancer (%)	**10 (19%)**		
Colorectal cancer (%)(1 x Colorectal liver metastasis*, 1 x Splenic flexure, 1 x Cecal)	3 (5%)		
Pancreatic cancer (%)	2 (3%)		
Small cell lung cancer (%)	1 (2%)		
Pseudomyxoma peritonei (%)	1 (2%)		
Metastatic bone disease (breast cancer recurrence) (%)	1 (2%)		
Anaplastic thyroid cancer (%)	1 (2%)		
Hepatocellular carcinoma (HCC)	1 (2%)		
Control (%)	**44 (81%)**		
Previous diagnosis of cancer (treatment completed) (%)	5 (9%)		

2.3. Patient and public involvement

2 former patients from Health and Care Research Wales' Public Involvement Community were involved in the research management of this work. They aided in the development of research protocols, patient information sheets and lay summaries as well as with interpretation of the practical relevance of results arising from the research ensuring that the research was relevant to patient and public need.

2.4. Sample handling and spectral acquisition

Serum samples were stored at − 80 °C until thawing for spectral acquisition. Raman spectra were obtained using an InVia Qontor (Renishaw, UK) Raman spectrometer. A previously designed high throughput system was used with inbuilt temperature control to prevent sample degradation [12]. In brief, the frozen serum was thawed then pipetted into a medical-grade stainless steel well plate at a volume of 200 μl per well. 3 repeat liquid serum spectra were obtained from each sample at the 785 nm laser line across a wavenumber range of 610–1720 cm^{-1}.

2.5. Spectral pre-processing

Data analysis within this study was carried out within the R programming environment [15]. Spectra were processed using Savitsky-Golay [16] filtering with a filter length of 9 and a polynomial order of 4 for spectral smoothing, fluorescent background has been corrected using a modified polynomial fitting algorithm [17] modeling with a 5th order polynomial, and spectra have been normalized using standard-normal-variate[18] followed by normalization to the phenylalanine peak at 1002 cm^{-1}. This procedure was developed in a high-performance computing workflow for optimization using a pre-processing package developed in-house [19]. A mean of 3 replicate spectra was taken as a patient's measurement.

2.6. Statistical analysis

Analysis is conducted in the programming language R using in-house scripts utilizing the base "stats" package containing several commonly used statistical techniques. Spectra from patients first are subjected to a Shapiro-Wilk test for normality to test appropriateness of using independent Welch t-test, and a Mann-Whitney U test for non-normality between the groups cancer and control. Each different type of cancer observed is then compared with the control set individually using a Wilcoxon signed rank. Welch's t-test, Mann-Whitney U test, and Wilcoxon signed rank remain robust for imbalanced data sets as in this study [20–22].

3. Results

126 patients were approached to take part, 5 patients declined to participate, and 121 patients consented (see Fig. 1 for STROBE diagram). Of these 58 patients provided a serum sample suitable for analysis. There were 22 females and 36 males with a median age of 59 years (range 28–90 years). Demographics of the cohort can be seen in Table 1.

Final diagnosis for the 58 patients was 10 cancers, 44 controls (no cancer or significant other disease diagnosis), and 4 significant other disease. For the purposes of this study the 4 significant other disease are not considered. Table 1 includes whether the patient has cardiovascular and/or hepatic disease, both of which have been shown to affect Raman spectra [38,39]. No patients in this cohort were noted to have a history of alcoholism or underlying infection.

We first compare the 2 groups cancer and control for each wavenumber. A Shapiro-Wilk test was first conducted to test for normality. Where peak differences could be seen, a Mann-Whitney U test (non-parametric) was conducted to test for significance between the collective cancer and control groups. The aim of this analysis is to identify peaks in the spectrum which are statistically significant. We then extend this to assess the significant peaks in individual cancers and controls using Wilcoxon signed rank for comparison.

Fig. 2 shows the mean control spectrum from 44 patients and the mean cancer spectrum from 10 patients in the RDC patient group, and the difference spectrum (mean control – mean cancer). The shaded purple and green regions outline the spectra ± the standard deviation. The spectra are produced by molecular vibrations within the serum sample, the peaks corresponding to vibrational modes. Subtle differences are observed between the cancer and control spectra with both increases and decreases in relative peak intensity. Statistical testing is applied to peak intensities to identify significance between cancers and

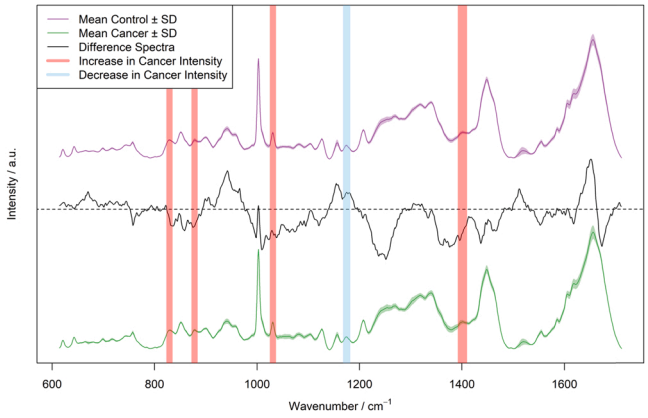

Fig. 2. Mean and standard deviation of Raman spectra for the cancer (n = 10) and control (n = 44) groups, and difference spectra between the two groups (control – cancer). Cancer spectra is offset for clarity on y-axis. Spectra are processed to corrected for background fluorescence, laser intensity fluctuations, and spectral noise. Peaks highlighted from Mann-Whitney U test. Red vertical regions indicate where a common increase in cancer intensity is observed again the control group, blue for decreases in cancer intensity.

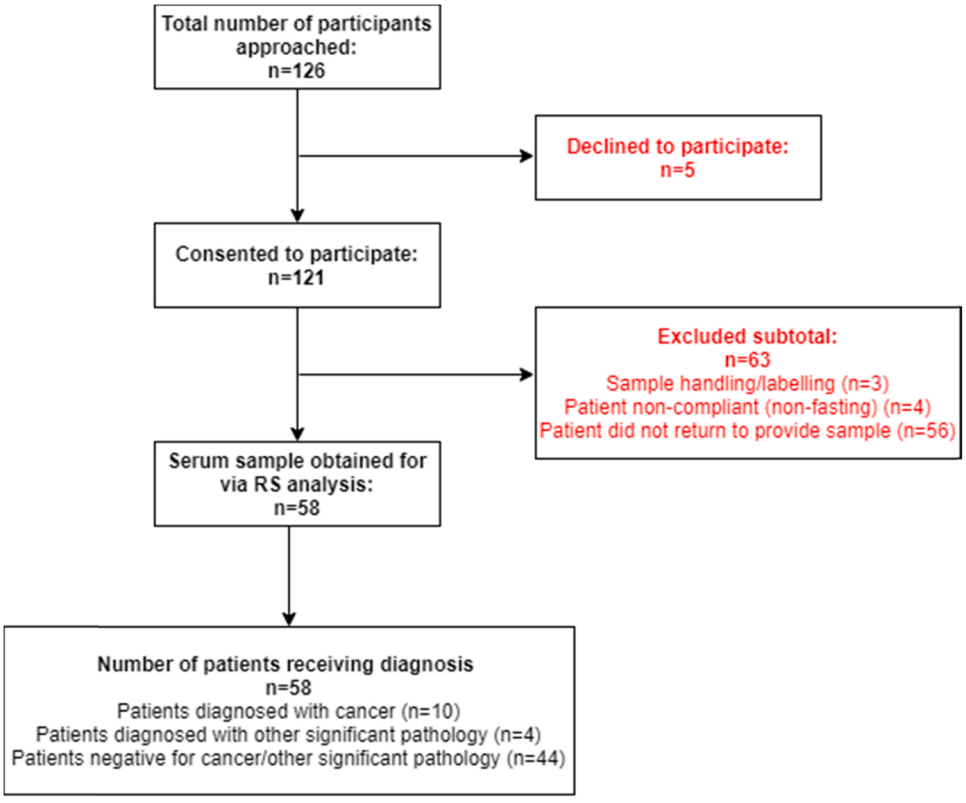

Fig. 1. STROBE diagram for the Raman-RDC study. High attrition due to patients' consent being obtained when not fasted and asked to return but did not.

controls. While the difference spectrum shows many regions where cancers and controls differ, only 5 are statistically significant due to large deviations in these other areas.

Several regions have been identified as significant between the control and cancer groupings; however, it must be noted these are cases where differences are in the same direction for each cancer i.e., either an increase in intensity compared to control, or a decrease overall. For 5 regions; 830 cm^{-1}, 878 cm^{-1}, 1031 cm^{-1}, 1174 cm^{-1}, and 1398 cm^{-1}. Tentative peak assignments are given in Table 2 in which a database of biological peak assignment has been produced from multiple literature sources and matched to significant peaks.

We now move to compare individual patients diagnosed with cancer in the RDC group with the control cohort. Due to small sample set this analysis is exploratory and is limited to direct comparison of cancer spectra to control group using one sample Wilcoxon ranked sign.

Table 3 shows the peaks where greater than 4 of the cancers observe significance from the control set with tentative Raman peak assignments included. In 3 peaks, we see all 10 cancers in the RDC set showing significance. These peaks at 878 cm^{-1}, 1449 cm^{-1} and 1519 cm^{-1}, tentatively attributed to proteins, amino acids, lipids, fatty acids, glycoproteins, carbohydrates, and carotenoids. Note this is increase and decreases in relative intensity, whereas important peaks described from the Mann-Whitney U test are where there are common increases and decreases between the cancers and controls on average.

Fig. 3 shows an example spectrum with each of the peaks in Table 3 show, along with the number of cancers where Wilcoxon sign rank shows significance between the cancer and control set.

4. Discussion

This study presents the first Raman diagnostic study in a cohort of non-specific symptomatic patients in an RDC setting. We found peaks in human blood sera spectra indicating significance between cancer and control samples in the cohort of 58 patients.

This study shows promising preliminary results in demonstrating a cohesive approach for cancer detection via a Raman based blood assay. Due to the nature of the study, single types of most cancers are seen, with only CRC and pancreatic cancer diagnosed in more than 1 patient. This limits analysis for some types to 1 vs. all controls. Future studies aim to collect larger datasets to improve confidence of cancer groupings

Table 2
Tentative peak assignment for significant wavenumbers between cancer and control globally. [23–27].

Peak position / cm^{-1}	Tentative assignment (Band)	Tentative component assignment	Cancer increase/ decrease	P value	Effect size
830	C-H ring stretching	Tyrosine; Phenylalanine; Mannose; Acetoacetate	↑	0.002 * *	Medium Effect
878		Glutamate; Mannitol; Unsaturated Fatty Acids; Coenzyme A, Carbohydrates; Arginine	↑	0.028 *	Small Effect
1031		Phenylalanine; Proline	↑	0.014 *	Medium Effect
1174	Amino Acid	Tyrosine; Histidine; Saturated Fatty Acids	↓	0.016 *	Small Effect
1398		Isoleucine; Glycine; Valine; Lysine; Acetoacetate; Carbohydrates	↑	0.048 *	Small Effect

Table 3
Tentative peak assignments for significant peaks identified individually between each cancer type and the control cohort. The following abbreviations are used; SCLC- Small cell lung cancer, SFC- splenic flexure carcinoma, PMP – pseudomyxoma peritonei, ATC – anaplastic thyroid cancer, CRLM - colorectal liver metastasis, Panc (1) – pancreatic cancer patient 1, Panc (2) – pancreatic cancer patient 2, HCC – hepatocellular carcinoma, MBD – metastatic bone disease (breast cancer recurrence), CC – cecal cancer [23–27].

Peak position / cm^{-1}	Tentative assignment (Band)	Tentative component assignment	Cancer difference seen (n patients with $p < 0.05$)
621	C-C twisting	Phenylalanine; Tyrosine	6 (SCLC, SFC, PMP, ATC, Panc (2), MBD)
643		Tyrosine; Proline	5 (SCLC, SFC, CRLM, HCC, MBD)
699		Cholesterol	9 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), Panc (2), HCC, MBD)
718		Tyrosine; Valine	9 (SCLC, SFC, PMP, ATC, CRLM, Panc (2), HCC, MBD, CC)
743		Tryptophan; Phosphoenolpyruvate; Threonine; Tyrosine; Riboflavin	8 (SCLC, SFC, PMP, ATC, Panc (1), Panc (2), HCC, CC)
757		Tryptophan; Valine	4 (SFC, PMP, ATC, Panc (1))
830	C-H ring stretching	Tyrosine, Phenylalanine, Mannose, Acetoacetate	9 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), HCC, MBD, CC)
851	C–C aliphatic stretch	Lysine; Valine; Tyrosine; Alanine; Isoleucine; Phenylalanine; Leucine; Histidine; Lactate	9 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), Panc (2), MBD, CC)
878		Glutamate; Unsaturated Fatty Acids; Coenzyme A, Carbohydrates; Arginine	10 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), Panc (2), HCC, MBD, CC)
941	ν(CC), $\nu(CCN)_{sym}$	Glutamate; Citric Acid, Aspartic Acid	8 (SCLC, SFC, PMP, ATC, CRLM, Panc (2), HCC, CC)
1031		Phenylalanine; Proline	8 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), HCC, CC)
1059		Histidine; Saturated Fatty Acids; Carbohydrates	7 (SCLC, SFC, PMP, CRLM, Panc (1), HCC, CC)
1082	C-O Vibration	Lactate, Triglyceride, Unsaturated Fatty Acids	8 (SCLC, SFC, PMP, CRLM, Panc (1), HCC, MBD, CC)
1127		Lactate; Valine; Saturated Fatty Acids; Carbohydrates	9 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), Panc (2), MBD, CC)
1157	(C-H)n, ν(CN)	Tyrosine; Glycine; Carotenoids	8 (SCLC, SFC, ATC, Panc (1), Panc (2), HCC, MBD, CC)
1174		Proline; Histidine; Methionine; Saturated Fatty Acids	9 (SFC, PMP, ATC, CRLM, Panc (1), Panc (2), HCC, MBD, CC)
1270		Histidine; Saturated Fatty Acids; Carotenoids	9 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), HCC, MBD, CC)
1319	δCH	Arginine; Glutamate; Histidine; Proline;	

(continued on next page)

Table 3 (*continued*)

Peak position / cm^{-1}	Tentative assignment (Band)	Tentative component assignment	Cancer difference seen (n patients with $p < 0.05$)
		Phosphoenolpyruvate; Histidine; Ascorbic Acid	7 (SCLC, PMP, CRLM, Panc (1), HCC, MBD, CC)
1340		Cystine; LysineThreonine; Glutamate	7 (SCLC, PMP, CRLM, Panc (1), HCC, MBD, CC)
1398		Isoleucine; Glycine; Valine; Lysine; Acetoacetate; Carbohydrates	7 (SCLC, SFC, ATC, CRLM, Panc (1), HCC, CC)
1449	δCH2, δCH3	Proteins; Phospholipids; Saturated Fatty Acids	10 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), Panc (2), HCC, MBD, CC)
1519		Glutamate; Carotenoids	10 (SCLC, SFC, PMP, ATC, CRLM, Panc (1), Panc (2), HCC, MBD, CC)
1587		Phenylalanine; Leucine; Serine	8 (SCLC, PMP, ATC, CRLM, Panc (1), HCC, MBD, CC)
1618		Tyrosine; Tryptophan	8 (SCLC, PMP, ATC, CRLM, Panc (1), HCC, MBD, CC)
1656		Amide I; Saturated Fatty Acids	9 (SCLC, PMP, ATC, CRLM, Panc (1), Panc (2), HCC, MBD, CC)

appearing. A larger sample set will also enable studies relating to clinically relevant but non-cancerous conditions. The current balance of confounding factors in the cancer/control patients in the study is reassuring for cardivascular, and only a small difference in hepatic issues. The future balance of extended trials is a factor that requires assessment and considerations to ensure that confounding factors do not skew or unintentionally define models.

RS from human blood sera from the RDC patient group has identified spectra differences between cancers and controls overall, notably in 4 peaks; 830 cm^{-1}, 878 cm^{-1}, 1031 cm^{-1}, and 1174 cm^{-1}, tentatively attributed to amino acids tyrosine, phenylalanine, glutamate, histidine, glycine, isoleucine, valine, lysine, and proline. In addition, these peaks can be attributed to unsaturated fatty acids and carbohydrates. These four regions see common collective increase or decrease on average for the cancer sera. Branched chain amino acids; valine, leucine, and isoleucine, can provide opportunistic energy sources for cells, which can feed into the altered metabolism of cancer cells [28]. Glutamate is mostly derived from glutamine, although can be synthesized from branched chain amino acids, which is an important link that is utilized by some tumors [29]. These markers may potentially be able to provide an immediate flag for further investigation for some cancer types via a threshold test. Confirmatory work will be undertaken to provide assurances of peak assignment using mass spectrometry.

Individual differences between cancer and control via statistical analysis identifies 3 peaks with significant for all 10 of the cancer patients. The peaks are at 878 cm^{-1}, 1449 cm^{-1} and 1519 cm^{-1}, tentatively attributed to proteins, amino acids, lipids, fatty acids, glycoproteins, carbohydrates, and carotenoids. These markers could potentially be used for preliminary testing of patients for changes which could be indicative of cancer presence, in this case a lower and upper threshold depending on the specific cancer type. This could be in the form of a percentage increase or decrease in peak intensity from controls with thresholds determined from a larger data cohort.

The hallmarks of cancer provide a logical framework for the biological processes and changes that tissues undergo for the development of tumorigenesis and ultimately, malignancy [30]. One of the hallmarks, reprogramming of energy metabolism, is a necessity for the fueling of uncontrolled cell proliferation and growth [30–32]. The biosynthetic, redox and bioenergetic demands of malignant cells must be satisfied via reprogramming the metabolic pathways and nutrient acquisition [28, 32]. The most notable pathways constituting on the energetical adaptations are: glutaminolysis, aerobic glycolysis, and pathological mitochondrial alterations [32]. Warburg observed one of the first distinguishing characteristics of cancer cell metabolism.[28] Irrespective of oxygen availability, cancer cells can sustain high rates of aerobic glycolysis for ATP generation [30,33,34]. The field of cancer metabolism expanded over years, revealing further metabolic alterations related to one carbon metabolism, including pentose phosphate pathway and tricarboxylic acid cycle (TCA cycle) with amino acids as alternative fuels [28,32].

In relation to the Warburg effect, 9 out of 10 of the cancers see statistical significance in the region 851 cm^{-1}, and for 8 out of 10 the region 1082 cm^{-1} both tentatively attributed to lactate. This could indicate the ability for Raman to detect changes in lactate uptake relating to cancer activity. We look to explore this effect further with increased sample numbers. In addition, several peaks in the spectrum with differences for multiple cancers can be attributed to components in the Krebs cycle such as glutamine, citric acid, coenzyme A, which can be affected in cancer development [35]. Differences are also seen with peaks tentatively attributed to arginine, of which arginine-derivatives (polyamines) are linked to cancer cell proliferation and aggressiveness

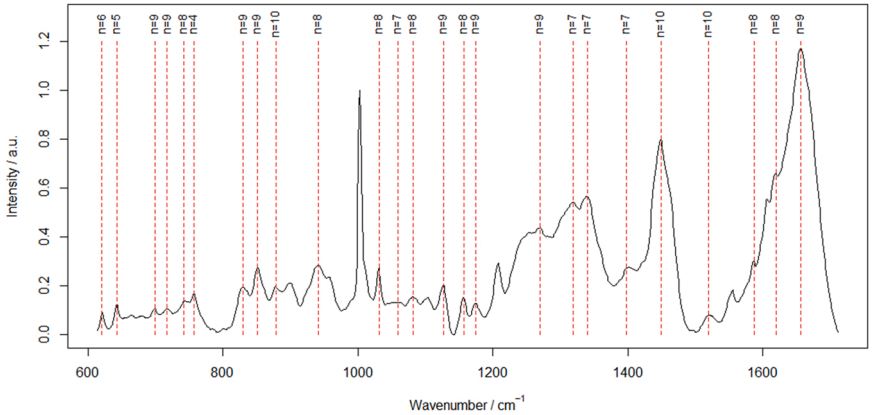

Fig. 3. Spectral plot displaying the regions with cancer peak intensities are statistically significant from control (red dashed line). Test performed using Wilcoxon signed rank between individual cancers and control set. The number fixed above the vertical peaks indicate the number of cancer patients in which statistical significance is observed. Tentative assignments can be found in Table 3.

[36,37].

Future work will verify the changes seen across cancers with mass spectrometry approaches to accurately assign metabolic species contributing to the tentative peak assignments. Additionally, development of artificial intelligence-assisted detection models for different cancer types through greater cohort size to include a diversity of cancer diagnoses is in progress the impact of this in an RDC could direct clinicians to targeted investigations saving money and time to diagnosis. With more data, confounding factors can be controlled for and included in diagnostic models which has not been possible with the small observational cohort in this study. Once test performance characteristics (sensitivity/ specificity/ area under curve) are known then formal clinical evaluation studies will be conducted with the potential for refining referral criteria for RDCs.

Data and Code availability

All original code has been deposited on GitHub and will be available at github.com/ferwoods/RDC following publication. Any additional information required to reanalyze the data reported in this paper is available from the corresponding author upon request.

CRediT authorship contribution statement

Study design and concepts were developed by FW, SC, DH and PD. DH and PD originated the funding for the study. FW and SC provided study data, conducted the literature review, and drafted the manuscript. FW analysed data and produced output analysis. FW, SC, NS, RH, AS, HG, HW, DH, and PD contributed to interpretation of results and finalizing of the manuscript.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:, PRD and DAH declare their involvement in CanSense Ltd, a recently incorporated cancer diagnosis spin-out company from Swansea University (company no: 11367637)., No other authors have any competing interests to declare.

Acknowledgments

This work was supported through Cancer Research Wales: Raman Spectroscopy and Colorectal Cancer: Transforming the USC Referral Pathway (Registered Charitable Incorporated Organization Number: 1167290). We are grateful to the RDC for facilitating access to their patients. We wish to acknowledge the input of Julie Hepburn and Ian Hills (of Health and Care Research Wales' Public Involvement Community) for their patient involvement activities.

References

[1] M. Astin, T. Griffin, R.D. Neal, P. Rose, W. Hamilton, The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review, Br. J. Gen. Pract. 61 (2011) e231–e243.
[2] M. Morris, E. Nolte, A. Aggarwal, J. Spicer, A. Purushotham, R. Sullivan, A. Aggarwal, C. Maringe, J. Spicer, M. Morris, A. Purushotham, E. Nolte, R. Sullivan, B. Rachet, The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study, Lancet Oncol. 21 (2020) 1023–1034, https://doi.org/10.1016/S1470-2045(20)30388-0.
[3] The NHS Long Term Plan, 2019. ⟨www.longtermplan.nhs.uk⟩ (accessed August 5, 2021).
[4] P. Vedsted, F. Olesen, A differentiated approach to referrals from general practice to support early cancer diagnosis – the Danish three-legged strategy, 2015. https://doi.org/10.1038/bjc.2015.44.
[5] B. Sewell, M. Jones, D. Fitzsimmons, H. Gray, C. Lloyd-Bennett, K. Beddow, M. Bevan, H. Wilkes, Rapid cancer diagnosis for patients with vague symptoms: a cost-effectiveness study, Br. J. Gen. Pract. 70 (2020) E186–E192, https://doi.org/10.3399/bjgp20×708077.
[6] C. Vasilakis, P. Forte, Setting up a rapid diagnostic clinic for patients with vague symptoms of cancer: a mixed method process evaluation study, BMC Health Serv. Res. 21 (2021) (2021) 1–11, https://doi.org/10.1186/S12913-021-06360-0.
[7] D. Chapman, V. Poirier, D. Vulkan, K. Fitzgerald, G. Rubin, W. Hamilton, S.W. Duffy, ARTICLE First results from five multidisciplinary diagnostic centre (MDC) projects for non-specific but concerning symptoms, possibly indicative of cancer, (n.d.). https://doi.org/10.1038/s41416–020-0947-y.
[8] Rapid Diagnostic Centres: Vision and 2019/20 Implementation Specification, (n.d.).
[9] B.E. C. K, E. P, R. AG, C. EJ, Top ten research priorities for detecting cancer early, The Lancet 4 (2019), e551, https://doi.org/10.1016/S2468-2667(19)30185-9.
[10] J.C. Martínez-Espinosa, J.L. González-Solís, M.L. Miranda-Beltrán, C. Soria-Fregoso, J. Medina-Valtierra, C. Frausto-Reyes, Detection of leukemia with blood samples using raman spectroscopy and multivariate analysis, In: Proceedings of the AIP Conference, 2009: pp. 99–103. https://doi.org/10.1063/1.3175637.
[11] J.L. González-Solís, J.C. Martínez-Espinosa, L.A. Torres-González, A. Aguilar-Lemarroy, L.F. Jave-Suárez, P. Palomares-Anda, Cervical cancer detection based on serum sample Raman spectroscopy, Lasers Med. Sci. 29 (2014) 979–985, https://doi.org/10.1007/s10103-013-1447-6.
[12] C.A. Jenkins, R.A. Jenkins, M.M. Pryse, K.A. Welsby, M. Jitsumura, C.A. Thornton, P.R. Dunstan, D.A. Harris, A high-throughput serum Raman spectroscopy platform and methodology for colorectal cancer diagnostics, The Analyst 143 (2018) 6014–6024, https://doi.org/10.1039/C8AN01323C.
[13] A. Sahu, S. Sawant, H. Mamgain, C.M. Krishna, Raman spectroscopy of serum: an exploratory study for detection of oral cancers, Analyst 138 (2013) 4161–4174, https://doi.org/10.1039/c3an00308f.
[14] J.L. Pichardo, O.B. Garcia, M.-R. Huerta-Franco, C.R. Alvarado, J.L. Pichardo-Molina, C. Frausto-Reyes, O. Barbosa-García, R. Huerta-Franco, J.L. González-Trujillo, C.A. Ramírez-Alvarado, G. Gutiérrez-Juárez, C. Medina-Gutiérrez, Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients, (2007). https://doi.org/10.1007/s10103-006-0432-8.
[15] R. Core Team, R: A Language and Environment for Statistical Computing, (2019). ⟨https://www.r-project.org/⟩.
[16] M.J.E. Savitzky, Abraham. and Golay, Smoothing and differentiation of data by simplified least squares procedures., 36, 1964. pp. 1627–1639.
[17] C.A. Lieber, A. Mahadevan-Jansen, Automated method for subtraction of fluorescence from biological Raman spectra, Appl. Spectrosc. 57 (2003) 1363–1367, https://doi.org/10.1366/000370203322554518.
[18] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra, Appl. Spectrosc. 43 (1989) 772–777, https://doi.org/10.1366/0003702894202201.
[19] F.E.R. Woods, P.R. Dunstan, PP. Raman: A package to handle pre-processing of Raman datasets in R, (2021). github.com/FERWoods/PPRaman (accessed April 20, 2021).
[20] B.L. Welch, The generalization of `student's' problem when several different population variances are involved, Biometrika 34 (1947) 28–35, https://doi.org/10.1093/biomet/34.1-2.28.
[21] H.B. Mann, D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, Annals Math. Stat. 18 (1947) 50–60, https://doi.org/10.1214/aoms/1177730491.
[22] M. Delacre, D. Lakens, C. Leys, Why psychologists should by default use Welch's t-test Instead of Student's t-test, Int. Rev. Soc. Psychol. 30 (2017) 92, https://doi.org/10.5334/irsp.82.
[23] J. de Gelder, K. de Gussem, P. Vandenabeele, L. Moens, Reference database of Raman spectra of biological molecules, J. Raman Spectrosc. 38 (2007) 1133–1147, https://doi.org/10.1002/jrs.1734.
[24] G. Zhu, X. Zhu, Q. Fan, X. Wan, Raman spectra of amino acids and their aqueous solutions, Spectrochim. Acta Part A: Mol. Biomol. Spectrosc. 78 (2011) 1187–1195, https://doi.org/10.1016/j.saa.2010.12.079.
[25] S. Giansante, H.E. Giana, A.B. Fernandes, L. Silveira, Analytical performance of Raman spectroscopy in assaying biochemical components in human serum, (n.d.). https://doi.org/10.1007/s10103-021-03247-8.
[26] N.C. Dingari, G.L. Horowitz, J.W. Kang, R.R. Dasari, I. Barman, Raman spectroscopy provides a powerful diagnostic tool for accurate determination of albumin glycation, PLOS One 7 (2012) 32406, https://doi.org/10.1371/journal.pone.0032406.
[27] C.G. Atkins, K. Buckley, M.W. Blades, R.F.B. Turner, Raman Spectroscopy of Blood and Blood Components, (n.d.). https://doi.org/10.1177/0003702816686593.
[28] E.L. Lieu, T. Nguyen, S. Rhyne, J. Kim, Amino acids in cancer, Experimental & Molecular Medicine. (n.d.). https://doi.org/10.1038/s12276-020-0375-3.
[29] B.-H. Choi, J.L. Coloff, The Diverse Functions of Non-Essential Amino Acids in Cancer, (n.d.). https://doi.org/10.3390/cancers11050675.
[30] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (2011) 646–674, https://doi.org/10.1016/j.cell.2011.02.013.
[31] J.R. Cantor, D.M. Sabatini, ' Affi, Cancer Cell Metabolism: One Hallmark, Many Faces, (2012) 881. https://doi.org/10.1158/2159-8290.CD-12-0345.
[32] R.J. Deberardinis, N.S. Chandel, Fundamentals of cancer metabolism Introduction and overarching principles, (n.d.). https://doi.org/10.1126/sciadv.1600200.
[33] P.P. Hsu, D.M. Sabatini, Cancer cell metabolism: Warburg and beyond, Cell 134 (2008) 703–707, https://doi.org/10.1016/j.cell.2008.08.021.
[34] O.S. G. A, Molecular classification and correlates in colorectal cancer, J. Mol. Diagn. 10 (2008) 13–27, https://doi.org/10.2353/JMOLDX.2008.070082.
[35] S. Cardaci, M.R. Ciriolo, TCA cycle defects and cancer: when metabolism tunes redox state, Int. J. Cell Biol. 2012 (2012), https://doi.org/10.1155/2012/161837.
[36] A.E. Pegg, Critical Review Mammalian Polyamine Metabolism and Function, (n.d.). https://doi.org/10.1002/iub.230.

[37] E.W. Gerner, F.L. Meyskens, Polyamines and cancer: old molecules, new understanding, Nat. Rev. Cancer 4 (2004) 781–792, https://doi.org/10.1038/nrc1454.
[38] K. Kochan, et al., Pathological changes in the biochemical profile of the liver in atherosclerosis and diabetes assessed by Raman spectroscopy, Analyst (2013), https://doi.org/10.1039/c3an00216k.
[39] A. Chaichi, A. Prasad, M.R. Gartia, Raman Spectroscopy and Microscopy Applications in Cardiovascular Diseases: From Molecules to Organs, Biosensors (2018), https://doi.org/10.3390/bios8040107.