Conference papers   Radiation and Environmental Science Centre

2023

Optimisation of Raman Spectral Processing for Classification of Radiotherapeutic Toxicity

Isha Behl

Dinesh Medipally

Chris Talbot

See next page for additional authors

Authors

Isha Behl, Dinesh Medipally, Chris Talbot, Aidan Meade, and Fiona Lyng

Optimisation of Raman spectral processing for classification of radiotherapeutic toxicity

Isha Behl[1, 2#], Dinesh K.R. Medipally[1, 2], Chris Talbot[3], Aidan D. Meade[1, 2*] and Fiona M. Lyng[1, 2*]

1. Centre for Radiation and Environmental Science, FOCAS Research Institute, Technological University Dublin, Ireland.
2. School of Physics, Clinical & Optometric sciences, Central Quad, Technological University Dublin, City Campus, Grangegorman, Ireland
3. Department of Genetics and Genome Biology, University of Leicester, UK

#Email – behlisha86@gmail.com, isha.behl@tudublin.ie

*Joint Senior Authors

ABSTRACT

Background:
Severe radiation toxicity can continue years after the completion of radiotherapy for prostate cancer patients. Currently, it is impossible to predict before treatment which patients will experience these long-term side effects. New approaches based on vibrational spectroscopy have advantages over lymphocyte and genomic assays in terms of minimal sample preparation, speed and cost. A high throughput method has been developed to measure Raman spectra from liquid plasma in a cover glass bottomed 96 well plate. However, the Raman spectra can show contributions from glass and water. The current study aims to optimise pre-processing steps to improve classification performance.

Methods:
Blood samples (n=32) were obtained from prostate cancer patients enrolled on the EU-funded REQUITE study (www.requite.eu) through a collaboration with the University of Leicester. Raman spectra were recorded from plasma samples using an in-house developed high throughput method. Extended multiplicative scattering correction (EMSC) was used for background correction.

Results:
Raman spectra of plasma were corrected individually and together for glass and water interference. It was observed that a good model efficiency was achieved for prostate cancer patients with no/minimal radiation toxicity (grade 0-1) and severe radiation toxicity (grade 2-3) after corrections with both glass and water rather than individually.

Conclusions:
It could be concluded from the study that to achieve good overall model efficiency, both glass and water corrections are required, when compared to no corrections, only glass corrections and only water corrections

Keywords: Raman spectroscopy, prostate cancer, plasma samples, glass, water, EMSC, radiation toxicity, partial least squares discriminant analysis

1. BACKGROUND

Prostate cancer ranks 4th worldwide as the most commonly diagnosed cancer and the 8th leading cause of cancer-related death worldwide [1]. In Ireland, according to the National Cancer Registry Annual Report 2022, prostate cancer makes up 1/3 of all invasive cancers and ranks 1st among the most common cancer in Ireland [2]. It kills ~600 Irish men per year and is responsible for 12% of all cancer death [2]. The risk category for an individual's prostate cancer can be predicted using the combination of tumour stage, Gleason score and the level of prostate-specific antigen (PSA) in the blood [3,4]. The standard treatment for high-risk localized prostate cancer is androgen deprivation therapy combined with radiotherapy (RT) [5].

RT is an important treatment modality for the management of cancer, with approximately 50% of all cancer patients receiving RT at some point in their treatment [6]. The lifetime risk of developing cancer for those born since 1960 is now estimated to be ~50% [7]. This means that RT will be required for 25% of the male population at some point in their lifetime. Despite recent significant technological advances to conform the dose of radiation to the tumour, normal tissue is always irradiated during radiotherapy and this can lead to the development of severe acute or late side effects for the patient [8]. Intrinsic radiosensitivity is a known cause of radiation toxicity and there is a large intrinsic patient-to-patient variability in response [9]. To date, no markers of tumour response to treatment or predictors of normal tissue toxicity are in clinical use; thus the present work addresses an area of urgent unmet clinical need.

Raman spectroscopy has advantages over cellular and -omic assays for biological and clinical sample characterisation in terms of requiring minimal sample preparation, speed and cost. Raman spectroscopy is based on the inelastic scattering of light by vibrating molecules and the positions, relative intensities and shapes of the bands in a Raman spectrum carry detailed information about the molecular composition of the sample. Raman spectra of cells and tissues are a superposition of contributions from each biochemical component, such as nucleic acids, proteins, lipids and carbohydrates, and can provide a rapid, label-free, non-destructive measurement of the biochemical fingerprint of biological material. Over the past 15 years, there have been numerous studies showing the potential of Raman spectroscopy for disease screening, diagnosis and prognosis and very promising results have been observed using cells, tissues and biofluids [10]. Recent studies, including those from our own laboratory, have shown the potential of Raman spectroscopy to characterise the radiation response of normal and tumour cells irradiated in vitro and of tumour tissue irradiated in vivo [11-14]. A recent pilot study in our laboratory investigated the potential of Raman spectroscopy to discriminate prostate cancer patients who showed severe radiation toxicity (Grade 2+ toxicity) from those who showed no/minimal toxicity (Grade 0-1 toxicity) [15]. In the latter study blood samples were collected from 42 patients on a Cancer Trials Ireland radiotherapy trial (formerly the All-Ireland Co-operative Oncology Research Group ICORG) 08-17 study, NCT00951535) and Raman spectra were recorded from peripheral blood lymphocytes. Classification models for toxicity were developed using Raman spectra of these samples together with known patient toxicity scores, with sensitivity and specificity rates of 90% and 85% achieved. The Raman analysis showed changes in the lipid profiles of the patient samples [15].

For the current study, Raman spectra were acquired from plasma from prostate cancer patients prior to RT using a cover glass bottomed 96 well plate as a substrate [16]. However, the spectra from plasma have an interference from glass and water, thus the current study aims to develop an optimised method for background correction of Raman spectra from liquid plasma.

2. METHOD AND MATERIALS

1.1 Sample collection and preparation

For the current study, the blood samples from no/minimal (grade 0 and 1) radiation toxicity (n=16) and from severe (grade 2 and 3) radiation toxicity (n =16) were obtained from prostate cancer patients before undergoing RT. These patients were enrolled on the EU-funded REQUITE study (www.requite.eu) through a collaboration with the University of Leicester. The patients were followed up post-RT, at 12 months and 24 months following radiotherapy and toxicity levels were recorded using the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) v4.0 grading system. For the current study, plasma samples were isolated from these blood samples by centrifugation at 3500 g for 5 min at 18°C. The samples were subsequently stored at −80°C prior to Raman acquisition.

1.2 Raman acquisition

Raman spectra were acquired using an in-house developed high throughput (HT)-Raman spectroscopy method [16] on a Horiba Jobin Yvon Labram HR800 UV micro-spectroscopy system (Horiba UK Ltd, Middlesex, UK). Briefly, 20 µl of liquid plasma was deposited on a cover glass bottomed 96 well plate (MatTek corporation) and Raman spectra of the plasma samples were acquired using a 785 nm laser focused through a 10× objective (N.A. 0.25). Spectra were recorded using a diffraction grating ruled with 300 lines/mm giving a spectral resolution of ~2.1 cm^{-1}. Spectra were recorded automatically from each well where the spectrometer was programmed using an in-house developed high throughput macro template. Each spectrum was acquired over the region from 400 cm^{-1} to 1800 cm^{-1}. Eleven spectra were recorded from each sample for each patient with a 20 s × 3 integration time. Multiple wavenumber calibration spectra of 1,4-Bis (2-methylstyryl) benzene were recorded along with each sample acquisition and used in spectral pre-processing.

1.3 Raman pre-processing

All spectral processing procedures were carried out within MATLAB (R2017a; Mathworks Inc., Natick, MA), along with in-house developed algorithms and procedures available within the PLS Toolbox (v 8.0.2, Eigenvector Research Inc., Wenatchee, MA).

The acquired Raman spectra were at first calibrated for wavenumber using an in-house standard of 1,4-Bis (2-methylstyryl) benzene and using in-house developed calibration procedures [14]. Extended multiplicative scattering correction (EMSC) was used for background corrections for glass and water corrections. Earlier this method was successfully applied to the correction of glass interference in the Raman spectra of oral exfoliated cells on a glass slide [17]. A similar, methodology has been implemented here to optimise background corrections for plasma samples.

According to the EMSC algorithm [18], a raw spectrum 'S' can be described as a linear superposition of the reference spectrum 'R' of a cell acquired on a substrate, a polynomial baseline signal 'B', and the background (interference) signal 'G':

$$S = R + G + B \quad\quad\quad\quad 4.1$$

The goal is to estimate the values of 'B' and 'G' such that they may be subtracted from the recorded spectrum. The algorithm weights the contribution of the reference background spectrum to the dataset. Then, it fits complete datasets to the reference spectrum to aid in accurate removal of background. EMSC also calculates a baseline signal from the reference spectrum of the cellular components and enforces it on the dataset. Ultimately, this leads to a baseline corrected dataset with background subtraction.

Consider a reference spectrum 'r' is obtained in a way that 'R' can be approximated by the product of reference spectrum and a certain weight:

$$R \approx C_r * r \quad\quad\quad\quad 4.2$$

where C_r is scalar for a given spectrum

Similarly, a spectrum from a background 'g' represent the spectral contribution of glass/water in the recorded plasma spectrum by 'G', as the product of the pure glass/water spectrum and a certain weight:

$$G \approx C_g * g \quad\quad\quad\quad 4.3$$

where C_g is scalar for a given spectrum, whereas, the slowly varying baseline 'B' can be represented using an appropriate 'Nth' order polynomial:

$$B_N = C_0 + C_1 x + C_2 x^2 + \cdots + C_N x^N \quad\quad\quad\quad 4.4$$

where 'N' is the order of the polynomial, and C_m for m = 0 → N represents the various coefficients in the polynomial.

So, the raw spectrum 'S', the reference spectrum 'r', the glass spectrum 'g', and the order of the polynomial 'N', all are the input for the EMSC algorithm, which returns estimates for C_r, C_g, and C_m for m = 0 → N.

$$S \approx [C_r * r] + [C_g * g] + \sum_{m=0}^{N} C_m x^m \quad\quad\quad\quad 4.5$$

The corrected spectrum can be expressed as,

$$T = \frac{S - [C_g * g] - \sum_{m=0}^{N} C_m \, x^m}{C_r}$$ ....⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯⋯...........................4.6

For carrying out EMSC, spectra from glass (n=10) and water (n=10) were acquired in a cover glass-bottomed 96 well plate, while as a pure biological component, spectra of plasma (n=10) were acquired on a calcium fluoride slide which has no interfering glass bands and so can be used as a reference spectrum. Then before subjecting it to the EMSC algorithm. glass, water and reference spectra were baselined and smoothed using a Savitzky–Golay filter (order = 5; window = 13). Then, this pre-processed background and raw patient plasma dataset were subjected to EMSC. The baseline corrections were part of EMSC. After carrying out EMSC, patient plasma dataset spectra were Savitzky–Golay filter (order = 5; window = 13) and standardised using vector normalisation before analysis.

1.4 Raman analysis

PLS-DA is a linear classification model based on partial least squares regression, where the y variable (the regression target) is encoded as the discrete spectral class (no/minimal or severe radiation toxicity) [19, 20]. PLS-DA aims to obtain maximum covariance between the independent and dependent variables of a multidimensional dataset by finding a linear subspace. This new subspace allows the prediction of dependent variables using a reduced number of factors, known as latent variables (LVs) [21]. The details of the PLS-DA approach used in this study are fully described previously [22].

3.   RESULTS AND DISCUSSION

It is known that glass has an interfering band in the region 1200-1400 cm^{-1} and water has an interfering band in the region 1600-1700 cm^{-1}[17]. Due to glass interference, all the bands in the region for proteins and lipids can be masked, while water interference can cause broadening of the Amide I protein band as shown in Figure 1a. In order to optimise the background corrections, the dataset was corrected with only glass, only water and both glass and water together as shown in Figure 1a, with the reference spectrum for EMSC acquired from a sample deposited on CaF$_2$. It was observed that in the case of only glass the broadening of 1600 -1700 cm^{-1} can be observed, while in the case of only water corrections, the background can be observed in the region 1200-1300 cm^{-1} (Figure 1a). Furthermore, it was observed with both glass and water corrections together no background effect can be observed on the mean spectrum.

Further, with the aim to differentiate no/minimal (grade 0 and 1) radiation toxicity (n=16) from severe (grade 2 and 3) radiation toxicity (n =16) and to understand the effects of various corrections PLS-DA models were developed. As shown in Figure 1b and Figure 1c, the score plot and loading (LV1) respectively of PLS-DA leave-one-patient-out cross-validation depict the effect of glass and water on the classification model. Further, background corrections were carried out with only glass, only water and both glass & water together. It can be seen in Table 1, with no corrections 77% sensitivity and 46% specificity were achieved. Good classification efficiency was achieved with no corrections mainly due to the background. With only glass and only water corrections, sensitivities of 92% & 80% respectively and specificities of 33% and 40% respectively were achieved (Table 1). Further, it was observed that better performance was achieved using both glass and water corrections together, with sensitivity of 86% and specificity of 44%. From the mean spectrum (Figure 1a) and model efficiencies (Table 1), it can be concluded that after correcting with only glass and only water, the background is hampering the overall efficiency of the models, whereas, with both glass and water corrections no background interference was evident. Therefore, through the current study, it can be concluded that it is recommended to carry out both glass and water corrections before analysing the plasma samples Raman spectra acquired on the cover glass bottomed 96 well plate.
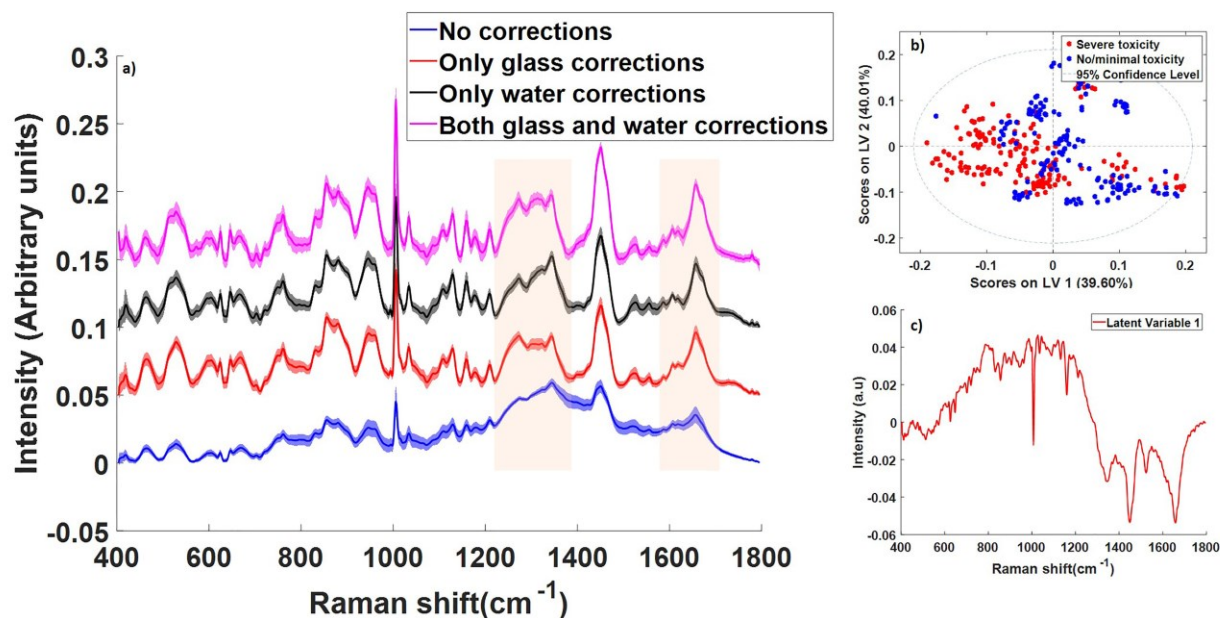
Figure 1 a) mean spectrum, b) PLS-DA leave-one-patient out with no corrections score plot, and c) Latent variable 1 with no corrections

Table 1: Comparative sensitivities, specificities and accuracy for various corrections

Corrections	Sensitivity (%)	Specificity (%)	Accuracy (%)
No	76	68	72
Glass	92	33	63
Water	80	40	60
Glass and water	86	44	65

ACKNOWLEGEMENT

REFERENCES

[1] Ferlay J, Colombet M, Soerjomataram I, *et al*. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources andmethods. *Int J Cancer*; 144: 1941–1953, 2019

[2] National Cancer Registry Ireland: Essential information on cancer in Ireland, November 2022: Prostrate cancer awareness month, https://www.ncri.ie/sites/ncri/files/infographics/Awareness%20Infographic%20Prostate%20Cancer%20A4.pdf [Accessed on 16/01/2023]

[3] Thompson IM, Ankerst DP, Chi C*, et al*. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ ml or lower. *JAMA*; 294: 66–70, 2005

[4] American Society of Clinical Oncology (ASCO): Prostate Cancer: Stages and Grades, https://www.cancer.net/cancer-types/prostate-cancer/stages-and-grades#:~:text=The%20Gleason%20scoring%20system%20is,cells%20receive%20a%20low%20score. 2021, [Accessed on 16/01/2023]

[5] Anderson EM and McBride SM, The Use of Androgen Deprivation Therapy in Combination With Radiation for Localized Prostate Cancer. Front. Urol. 2:890814, 2022, doi: 10.3389/fruro.2022.890814

[6] Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. Int J Med Sci. 2012;9(3):193-9. doi: 10.7150/ijms.3635, 2012

[7] Ahmad AS, Ormiston-Smith N, Sasieni PD. Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. Br J Cancer. 2015 Mar 3;112(5):943-7. doi: 10.1038/bjc.2014.606, 2015

[8] Barnett GC, West CM, Dunning AM, Elliott RM, Coles CE, Pharoah PD, Burnet NG. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. Nat Rev Cancer. 2009 Feb;9(2):134-42. doi: 10.1038/nrc2587, 2009

[9] Ohri N, Dicker AP, Showalter TN. Late toxicity rates following definitive radiotherapy for prostate cancer. Can J Urol, 19(4):6373-80, 2012

[10] K. Kong et al. / Advanced Drug Delivery Reviews 89. 121–134, 2015

[11] Matthews Q, Brolo A, Lum J, Duan X, Jirasek A. Raman spectroscopy of single human tumour cells exposed to ionizing radiation in vitro. Phys Med Biol.;56(1):19-38, 2011 doi: 10.1088/0031-9155/56/1/002.

[12] Maguire, A., Vegacarrascal, I. & White, L.. Analyses of ionizing radiation effects in – vitro in peripheral blood 1 lymphocytes with Raman spectroscopy. *Radiation Research*, vol. 183, no. 4, pp. 407-416, 2015 doi.org/10.1667/RR13891.1

[13] Harder, S., Isabelle, M., DeVorkin, L. *et al.* Raman spectroscopy identifies radiation response in human non-small cell lung cancer xenografts. *Sci Rep* **6**, 21006, 2016 https://doi.org/10.1038/srep21006

[14] Meade AD, Maguire A, Bryant J*, et al*. Prediction of DNA damage and G2 chromosomal radio-sensitivity ex vivo in peripheral blood mononuclear cells with label-free Raman microspectroscopy. *Int J Radiat Biol*; 95: 44–53, 2018

[15] Cullen D, Bryant J,Maguire A, et al. Raman spectroscopy oflymphocytes for the identification of prostatecancer patients with late radiation toxicityfollowing radiotherapy.Translational Biophotonics.;2:e201900035, 2020 https://doi.org/10.1002/tbio.2019

[16] Medipally DKR, Maguire A, Bryant J*, et al*. Development of a high throughput (HT) Raman spectroscopy method for rapid screening of liquid blood plasma from prostate cancer patients. *Analyst*; 142: 1216–1226, 2017

[17] Behl I, Calado G, Malkin A, et al. A pilot study for early detection of oral premalignant diseases using oral cytology and Raman micro-spectroscopy: Assessment of confounding factors. J Biophotonics.;13(11):1-12. doi:10.1002/jbio.202000079, 2020

[18] Kerr L.T., Hennelly B.M., A multivariate statistical investigation of background subtraction algorithms for Raman spectra of cytology samples recorded on glass slides, Chemometrics and Intelligent Laboratory Systems,158, 61-68, 2016

[19] Wold S, Sjöström M and Eriksson L. PLSregression: a basic tool of chemometrics. *Chemometr Intell Lab Syst*; 58: 109–130, 2001

[20] Brereton RG and Lloyd GR. Partial least squares discriminant analysis: taking the magic away. *J Chemom*; 28: 213–225, 2014

[21] Gromski PS, Muhamadali H, Ellis DI*, et al*. A tutorial review: metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta*; 879: 10–23, 2015

[22] Dinesh K. R. Medipally, Daniel Cullen, Valérie Untereiner, Ganesh D. Sockalingum, Adrian Maguire, Thi Nguyet Que Nguyen, Jane Bryant, Emma Noone, Shirley Bradshaw, Marie Finn, Mary Dunne, Aoife M. Shannon, John Armstrong, Aidan D. Meade, and Fiona M Lyng Therapeutic Advances in Medical Oncology 10.1177/1758835920918499, 2020