2018

Prediction of DNA Damage and G2 Chromosomal Radio-Sensitivity Ex-vivo in Peripheral Blood Mononuclear Cells with Label-Free Raman Microspectroscopy

Aidan Meade
Technological University Dublin, aidan.meade@dit.ie

Adrian Maguire
Dublin Institute of Technology, adrian.maguire@mydit.ie

Jane Bryant
Dublin Institute of Technology, jane.bryant@dit.ie

Daniel Cullen
Dublin Institute of Technology

Dinesh Medipally
Dublin Institute of Technology

See next page for additional authors

Recommended Citation

Meade, A., Maguire, A. & Bryant, J. (2018). Prediction of DNA Damage and G2 Chromosomal Radio-Sensitivity Ex-vivo in Peripheral Blood Mononuclear Cells with Label-Free Raman Microspectroscopy. *International Journal of Radiation Biology*, March 29, pg. 1-10. doi.org/10.1080/09553002.2018.1451006

Authors
Aidan Meade, Adrian Maguire, Jane Bryant, Daniel Cullen, Dinesh Medipally, Lisa White, Brendan McClean, Laura Shields, John Armstrong, Mary Dunne, Emma Noone, Shirley Bradshaw, Marie Finn, Aoife M. Shannon, Orla Howe, and Fiona Lyng

Prediction of DNA damage and G2 chromosomal radio-sensitivity ex-vivo in peripheral blood mononuclear cells with label-free Raman micro-spectroscopy

Aidan D. Meade[*1, 2], *Adrian Maguire*[1, 2], *Jane Bryant*[2], *Daniel Cullen*[2,3], *Dinesh Medipally*[2,3], *Lisa White*[2,3], *Brendan McClean*[4], *Laura Shields*[4], *John Armstrong*[5,6], *Mary Dunne*[5], *Emma Noone*[5], *Shirley Bradshaw*[5], *Marie Finn*[5], *Aoife M. Shannon*[6], *Orla Howe*[2,3], *Fiona M. Lyng* [1, 2]

1. School of Physics, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland.
2. DIT Centre for Radiation and Environmental Science, Focas Research Institute, Dublin Institute of Technology, Camden Row, Dublin 8, Ireland.
3. School of Biological Sciences, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland
4. Department of Medical Physics, Saint Luke's Radiation Oncology Network, St Luke's Hospital, Dublin, Ireland
5. Department of Radiation Oncology, Saint Luke's Radiation Oncology Network, St Luke's Hospital, Dublin, Ireland.
6. Cancer Trials Ireland, Dublin, Ireland

Prediction of DNA damage and G2 chromosomal radio-sensitivity *ex-vivo* in peripheral blood mononuclear cells with label-free Raman micro-spectroscopy

Aidan D. Meade[*1,2], *Adrian Maguire*[1,2], *Jane Bryant*[2], *Daniel Cullen*[2,3], *Dinesh Medipally*[2,3], *Lisa White*[2,3], *Brendan McClean*[4], *Laura Shields*[4], *John Armstrong*[5,6], *Mary Dunne*[5], *Emma Noone*[5], *Shirley Bradshaw*[5], *Marie Finn*[5], *Aoife M. Shannon*[6], *Orla Howe*[2,3], *Fiona M. Lyng* [1,2]

1. School of Physics, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland.
2. DIT Centre for Radiation and Environmental Science, Focas Research Institute, Dublin Institute of Technology, Camden Row, Dublin 8, Ireland.
3. School of Biological Sciences, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland
4. Department of Medical Physics, Saint Luke's Radiation Oncology Network, St Luke's Hospital, Dublin, Ireland
5. Department of Radiation Oncology, Saint Luke's Radiation Oncology Network, St Luke's Hospital, Dublin, Ireland.
6. Cancer Trials Ireland, Dublin, Ireland

Abstract

Purpose

Liquid biopsies are a potentially rich store of biochemical information that can be linked to an individual's response to therapeutic treatments, including radiotherapy, and which may ultimately play a role in the individualization of treatment regimens. Peripheral blood mononuclear cells (PBMCs) can be used for the biochemical profiling of the individual, but also, being living cells, can provide insights into the individuals response to ionizing radiation exposure.

Material and Methods

The present study attempts to link the biochemical profile of lymphocytes within PBMCs obtained through Raman spectroscopy, to in-vitro measures of low-dose (< 0.5Gy) DNA damage response and cytogenetic metrics of radiosensitivity in a cohort of healthy controls and prostate cancer patients (from CTRIAL-IE(ICORG) 08-17, NCT00951535). All parallel metrics to the Raman spectra of the cells were obtained ex-vivo in cycling peripheral blood lymphocytes, with radiosensitivity estimated using the

G2 chromosomal assay and DNA damage assessed using γH2AX fluorescence. Spectra from a total of 26 healthy volunteers and 22 prostate cancer patients were obtained.

Results

The links between both measures of cellular response to ionizing radiation and the Raman spectra were modelled using partial least squares regression (PLSR) and support-vector regression (SVR). It was found that neither regression approach could predict radiation-induced G2 score well, but could predict γH2AX MFI with the SVR outperforming PLSR, implying a non-linear relationship between spectral measurements and measures of DNA damage.

Conclusion

Raman spectroscopy of PBMCs represents a label-free approach for prediction of DNA damage levels for either prospective or retrospective analysis.

Keywords: Raman spectroscopy, individual radiation sensitivity, DNA damage, γH2AX fluorescence, G2 radiosensitivity, partial least squares regression, support vector regression.

1. Introduction

Individualisation of therapeutic strategy for radiotherapy patients remains a goal of clinical predictive testing prior to treatment regimes. This response is dependent largely on various innate genetic, epigenetic and lifestyle characteristics that have yet to be fully characterised. Studies of inherited conditions leading to extreme radiosensitivity, such as Ataxia Telangiectasia (AT), Nijmegen Breakage Syndrome (NBS1) and Fanconi's Anemia (FA) (1) have led to an understanding of the genetic signals which are involved in the response to ionising radiation exposure, and in all mammalian cells (2). One of the key drivers of radiosensitivity is that genetic mediators of the DNA

damage response (DDR) such as ATM in AT, NBS1 (from MRN complex) in NBS, BRCA2 in Fanconi's, are mutated and cause a disruption in the DDR leading to these clinical genetic conditions(3,4). However a distribution in intrinsic radiosensitivity exists in humans, which in the general population follows a Gaussian distribution, encompassing radioresistant, normal and radiosensitive individuals (5,6). Although the molecular mechanisms in radiobiological response are relatively well characterised, identification of individuals at risk of adverse therapeutic response remains a challenge requiring the integration of genomic analyses with other functional assays (7). In the present study the potential for Raman spectroscopy to measure DNA damage and intrinsic radiosensitivity *ex-vivo* using peripheral blood mononuclear cells (PBMCs) is examined.

Raman spectroscopy is well characterised as a methodology allowing the objective classification of cancerous tissue pathology (8,9), and even the identification of pre-cancerous staging (10). It has the advantage of not requiring any additional biochemical labelling, and being applicable to solid and liquid biopsies (11), and even *in-vivo* (12). In addition the use of this analytical technique for the monitoring and characterisation of radiobiological responses is an important step in developing its applications within the therapeutic sphere. Early studies here demonstrated the ability to discriminate irradiated cells from controls at high (>2Gy) doses (13). More recent studies have confirmed the potential of Raman spectroscopy, and Fourier Transform Infrared spectroscopy, to discriminate dose dependent responses to both targeted and non-targeted radiation damage at doses as low as 5mGy (14,15), and importantly to allow the discrimination of radiobiological effects at low doses (<0.5Gy) in PBMCs (16). In this latter study it was suggested that spectral signatures which correlate with DNA damage levels, as

measured using γH2AX fluorescence intensity are present in Raman spectra of PBMCs (16). As γH2AX phosphorylation occurs after the sensing of DNA double strand breaks (DSB) by the ataxia telangiectasia mutated gene (ATM) in the DNA Damage Response, the γH2AX assay has been widely used as an assay for the measurement of DSB sensing and repair (17,18). However, while the assay is a good measure of both DNA damage and damage sensing of DSBs, its use in dosimetry and assessment of individual radiosensitivity has been hindered by the large variation in individual baseline levels of H2AX phosphorylation (19-21).

Since the γH2AX assay measures the levels of sensed DNA damage post-ionising radiation exposure it is related to the function of the ATM gene and as such has some relationship to intrinsic radiosensitivity. The G2 chromosomal assay is an in-vitro predictive assay of patient intrinsic radiosensitivity, that produces radiation-induced de-novo chromatid aberrations in lymphocytes in the G2 phase of the cell cycle before entering mitosis (22). The assay is a reliable predictor of radiosensitivity and correlates well with cancer predisposition (23). Increased aberration yields have been confirmed in a wide range of cancer types relative to control cohorts including head and neck cancer (24), retinoblastoma (25), melanoma (26), and breast cancer (27-34). Similarly the assay displays higher aberration yields in groups suffering from syndromes exhibiting radiosensitivity including Bloom's Syndrome, Fanconi's Anaemia (35,36), and Ataxia Telangiectasia (37). While the assay demonstrates good levels of reproducibility between operators and laboratories (36,38), it requires 3- 5 days to complete.

In this study, Raman spectroscopy was performed on the lymphocyte component of PBMCs extracted from a group of prostate cancer patients and healthy control volunteers. In parallel, measurements of γH2AX DSB foci and chromosomal

aberrations using the G2 chromosomal assay were performed 1-hour after in-vitro irradiation. Partial least squares regression and support vector regression of spectra onto both the radiation-induced G2 radiosensitivity scores (riG2) and γH2AX mean fluorescence intensity (MFI) were performed to determine whether robust spectral signatures of DNA damage and intrinsic radiosensitivity exist within Raman spectra.

2. Methods

2.1 Cohort Characteristics and Ethical Approval

The objective of the present study is to elucidate whether Raman spectroscopy can be used as a valid alternative to metaphase spreads or staining protocols for the measurement of DNA damage and/or radiosensitivity. Therefore the control cohort (non-age matched) and patient cohort are structured here to provide a wide range of radiosensitivities to justify the conclusions of the work.

Ethical approval was awarded by the Dublin Institute of Technology Ethics Committee (2012) for the collection of PBMCs from blood donations from a total of 26 volunteers (with an age range from 21-26) for this study. Volunteers were from both sexes and comprised both smokers and non-smokers.

Ethical approval was separately awarded by the Dublin Institute of Technology Ethics Committee (2012) for the collection of blood donations from prostate cancer patients from St. Luke's Hospital, St. James's Hospital, Beacon Hospital, and Beaumont Hospital (all in Dublin) for the purposes of the study. The study was also approved by Cancer Trials Ireland (formerly the All-Ireland Co-operative Oncology Research Group (ICORG)) as a translational sub-study to CTRIAL-IE (ICORG) study 08-17 (ICTRP ID:

NCT00951535; “A Prospective Phase II Dose Escalation Study Using IMRT for High Risk N0 M0 Prostate Cancer”) with an amendment to existing ethical approval from St. Luke’s Hospital, and St. Luke’s Radiation Oncology Units at St. James’s Hospital and Beaumont Hospitals, Dublin. Patients enrolled in the study were male prostate cancer patients who were consented prior to treatment (therefore none of the Raman and/or biological data presented here was taken from the patients whilst in their treatment phase (i.e. hormone therapy or radiotherapy), but rather were taken at baseline post-diagnosis), and were both smokers and non-smokers within the age range from 58 to 85. Spectra were acquired from the PBMC fraction of their blood for each of 22 different patients prior to radiotherapy.

2.2 Cell culture and PBMC isolation from peripheral blood

Procedures for isolation of PBMCs from peripheral blood have been reported previously (16). Briefly, after drawing of fresh blood into Lithium-heparin tubes, PBMCs were isolated within 24 hours of collection. After isolation of the PBMC layer from whole blood using Histopaque and gentle centrifugation (400g) the PBMC fraction was washed 3 times in de-ionised phosphate buffered saline (DPBS) and pelleted (through centrifugation at 250g). The pellet was re-suspended in RPMI with all supplements, divided into three flasks containing medium, and incubated for 72 hours at 37°C, 5% CO_2 with 2.5%(v/v) phytohaemagglutinin.

2.3 Sample prep and Raman spectroscopy

Cells in suspension were first centrifuged at 250g for 5 minutes. Supernatant was removed and the cells were fixed using 200µl of 2% paraformaldehyde in PBS. From the suspension, 40µl was drop cast onto calcium fluoride (CaF_2) slides. The remaining suspension was centrifuged, paraformaldehyde was removed and cells were then stored in ethanol at -20ºC for γ-H2AX staining at a later date. Paraformaldehyde was removed

from the slides for Raman spectroscopy and they were then rinsed in deionised H_2O for 5 minutes. Washing was performed three times and then the slides were allowed to dry for Raman spectroscopic measurements.

Spectroscopic measurements were performed using a 660nm excitation on a LabRam HR800 (Horiba UK Ltd.) system with all calibration and measurement procedures as indicated previously (16). Spectra were recorded from 40 lymphocytes for each sample (dose and donor) using a rastering approach such that each spectrum represents the spectrum from the whole cell. All spectral processing procedures were conducted as indicated previously (16) using 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).

2.4 Whole blood irradiation

The whole blood samples were irradiated using a 6MV photon beam produced by an Elekta Precise linear accelerator (LINAC) at St. Luke's Hospital, Dublin, operating at a nominal dose rate of 6Gy/min . The LINAC was calibrated in accordance with the 1990 IPSM code of practice by the Medical Physics Department at St. Luke's Hospital(39), such that 100 Monitor Units (MU, a measure of "beam on" time) delivered a dose of 1Gy at 1.4 cm deep in water positioned 100cm from the source for a 10x10 cm^2 field. In order to achieve a uniform irradiation of flasks in practice, the irradiation conditions were altered from those at calibration. A 30x35cm^2 field was used to deliver each dose. The flasks were also positioned 10cm deep in a water equivalent phantom 90cm from the source. At 90cm from the source 100MU delivers a dose of 0.812Gy at 10cm deep in water for a 10x10cm^2 field. The number of MU required to deliver each of the doses outlined above must be corrected for the different scatter conditions present with the larger field size (30x35cm^2). A correction factor of 1.1372 was therefore applied, which

is the ratio of the field area of a large field to a smaller one. Thus at 90cm from the source, 100MU delivers a dose of 0.9234 Gy (0.812×1.1372), and so the delivery of 0.05Gy required 6 MU and 0.5Gy required 55 MU (MU were rounded up to the nearest whole number as partial MU could not be delivered on the LINAC).

The calculated doses were verified using Gafchromic EBT3 film (Ashland Inc., NJ). The film was calibrated against a Farmer type ionization chamber using the triple channel dosimetry method(40). The film was scanned using the single scan protocol(40) on an Epsont Expression 10000 XL scanner using the recommended scanning resolution of 72 dpi in a 48-bit RGB format (40-42). Glass was placed over the calibration and test film during scanning to minimize ringing artifacts. The film was analyzed using FilmQA Pro (Ashland Inc.).

2.5 G2 chromosomal aberration assay

The G2 chromosomal radiosensitivity assay was performed to produce a G2 radiosensitivity score per patient sample and this was used as a parallel reference measurement of radiation damage and sensitivity for a comparison to Raman spectroscopic measurements. The procedure undertaken to perform the assay is adapted from earlier work (43) and is described briefly below.

Thirty minutes after irradiation 0.2 ml of colcemid (10μg/ml) (Roche Diagnostics GmbH) was added to *in-vitro* cultured whole blood cells and incubated at 30C for 60 minutes to arrest the lymphocyte cells in metaphase. Lymphocytes were then isolated from other whole blood cells through centrifugation at 1600rpm for 10 minutes. After centrifugation, the lymphocyte cells were plunged in ice to cool rapidly, supernatant was removed and replaced with 10ml of pre-cooled 0.075M KCl hypotonic and placed in ice for 20 minutes to lyse the cells. After centrifugation at 1600rpm for 10 minutes, cells were fixed twice with 3:1 methanol:glacial acetic acid, and stored at 4^{o}C

overnight or for longer if required. Slides were prepared by rinsing in methanol 24 hours prior to use, and then briefly washed with de-ionised water. One-to-two drops of the cell suspension were dropped onto the slides. Three-percent Geimsa (GURR) in pH 6.8 buffer was added to the slides for 15 minutes to stain the DNA in the lymphocyte cells. Slides were washed in pH 6.8 buffer and left to dry before being mounted in DPX. Slides were left for 24 hours (minimum) before analysis. The number of chromosomal aberrations was recorded for 50 cells per slide and was multiplied by two to get a percentage of aberrations for each slide. This constituted the G2 radiosensitivity score and was recorded for each dose (0, 50 and 500mGy) per patient sample. A radiation-induced G2 score (riG2) was calculated by subtracting the sham-irradiated (0 mGy) G2 score from the irradiated G2 scores (50 and 500mGy) separately.

Previous studies have employed a scoring protocol for 100 metaphases but only a 500mGy dose point, and as the present study includes a 50mGy dose point, this scoring approach was deemed reasonable to maintain workflow.

2.6 Measurement of levels of DSB foci using γ-H2AX fluorescence

Methods for measurement of γH2AX foci were described in detail previously (16). Briefly, after preparation of cells for Raman spectroscopy, lymphocytes for the γH2AX assay were stored at -20^{o}C until the time of analysis and were then permeabilised via incubation for 5 minutes at room temperature in 200-1000µl of 0.25% (v/v) Triton X-100 in PBS, followed by resuspension in 200µl of blocking solution (PBS containing 2% (w/v) BSA) and incubated for 30 minutes at room temperature. After removal of the blocking solution the cells were resuspended in 150µl of primary antibody solution (Anti-phospho-histone H2AX) for 2 hours incubated at room temperature and after washing cells in PBD, 150µl of secondary antibody solution (Alexa Fluor 488) was

added. The samples were incubated for 1 hour in the dark and subjected to a final washing step before fluorescence was measured using was a BD Accuri C6 flow cytometer, with signals from debris and aggregates removed using forward and side scatter characteristics. From a minimum of 10000 events per sample the γH2AX MFI was calculated.

2.7 Development of regression models for prediction of γ-H2AX fluorescence and G2 chromosomal radiosensitivity from Raman spectra

2.7.1. Partial Least Squares Regression

Partial least squares regression (PLSR) is an analytical technique that is used extensively in spectroscopy to regress, using a linear model, a set of physical measurements (in spectroscopy the physical measurements are the intensities at each wavenumber or wavelength) onto a vector or matrix of target agents or concentrations (14). PLSR is a linear modelling approach that relates the target analyte (Y) to a matrix of spectra (X), and a matrix of residual errors (E) as (44):

$$Y = XB + E \qquad \text{Equation 1}$$

where B is a matrix of regression coefficients. In order to produce parsimonious models which use a minimum number of variables, both X and Y undergo dimensionality reduction using principal components analysis as:

$$X = TP^T + E_x \qquad \text{Equation 2}$$

and

$$Y = UQ^T + E_Y \qquad \text{Equation 3}$$

where P and Q are matrices of latent variables, T and U are matrices of scores and E_X and E_Y are matrices containing residuals. The matrices of scores of latent variables T and U can then be related through a linear regression as follows:

$$U = TD + H \qquad \text{Equation 4}$$

where D is a diagonal matrix of regression parameters and H is the residual matrix. In general, U is related to T by only a subset of the matrix of scores and latent variables.

2.7.2. Support Vector Regression

Support vector regression (SVR) is a statistical learning approach employing support vector machines for the purposes of a regression task. A detailed discussion of their properties is beyond the scope of the current study, and the reader is referred to the seminal references in this section.

Support vector machines (SVMs), in a classification problem, are capable of learning the position of any complex hyperplane separating two or more classes (45), and within a regression task, they can similarly learn any complex function linking a matrix of X data to Y targets (46). The advantage of the use of SVR in their application to high-dimensional problems such as the present one, is their insensitivity to outliers or data noise. Briefly, SVR is an application of statistical learning for non-linear regression of multivariate or multidimensional data, where a function *f(x)* is learnt on the basis of the input patterns, *x*, using a set of weights *w* such that:

$$f(x) = \langle w, x \rangle + b \qquad \text{Equation 5}$$

where *b* is a real constant and $\langle w, x \rangle$ is the dot product between *w* and *x*. The 'support vectors' are then the input patterns which define *f(x)*, and the number of support vectors defines the model complexity, which naturally should be minimized. To allow the algorithm to be insensitive to noise the SVM introduces the concept of a 'soft-margin' or error such that 'slack-variables' are allowed to influence deviations from the function by a degree specified by ξ_i and ξ_i^* subject to retaining a precision of ε (epsilon) within the regression. To prevent large deviations from the function two approaches are then used: (i) the introduction of a *penalty* parameter *C* to minimize the influence of the slack variables (C-SVR) or (ii) the introduction of a parameter *nu*, *ν*, specifying the upper limit on the fraction of margin errors allowed and the lower limit on the number of support vectors required in the regression (ν-SVR) (47). In the present study the ν-SVR approach was used.

The final property of SVMs which must be introduced in the context of SVR is the *kernel trick*. In the formulation defined in equation 5, the algorithm is effectively a linear regression. In order to allow non-linear regression the kernel trick is employed which linearizes the input patterns by mapping to a non-linear feature space where the linear function *f(x)* may be learnt. Various kernel tricks are available, including linear, polynomial and sigmoidal. In the present study a Gaussian radial-basis function (RBF) kernel was employed as a non-linear mapping, whose width was denoted by a parameter γ, *gamma*. This kernel represents a mapping which is adaptable to any non-linear function, as compared to the other options which are detailed above, and therefore does not require knowledge of the order of the non-linear regression function prior to the analysis.

2.7.3. Model Training and Testing

For regression of the spectral data against each of γH2AX MFI and riG2 targets a database of spectra from 26 controls and 22 patients at each of the three dose points used (0Gy, 0.05Gy and 0.5Gy) was created, with these matched to the relevant measurements of each biological target. Each spectral subset (i.e. for an individual donor or patient and radiation dose) was considered unique within the current analysis such that a total of approximately 3×48=144 subsets each containing 40 spectra of PBMCs were created. Within each spectral subset the spectra were averaged by randomly selecting groups of 10 such that spectral noise could be minimized (14). A set of training data was then created by stratified-random selection of 70% of the spectral subsets within each quartile of the reference γH2AX MFI or riG2 score data, such that the algorithms could be trained with spectral data that was matched to biological targets over the full range available. The remaining spectral subsets (30%) of the data were used to test each models generalizability to unseen data, representing the likely performance of the algorithm in the real world. A pooled analysis of all of the data from controls and patients in this way presents the algorithms with spectral data from individuals with a range of radiosensitivities, susceptibility to damage, and propensity for repair.

Optimization of the number of latent variables (or complexity) within the PLSR models was performed using five-fold cross-validation to find the minimum number of latent variables (LVs) that minimized root-mean-squared error of regression at calibration (RMSEC) on the training set. Similarly the SVR models were tuned by varying the values of the C, γ and ν parameters on the training matrix to find the optimal set of parameters for regression. For both models their generalizability to unseen data was measured using the root-mean squared error of prediction on the testing set (RMSEP).

For both modelling approaches the spectral averaging steps, the random sorting of spectra between training and testing sets, and model optimization steps were performed on 20 separate occasions to obtain robust measures of the performance of each regression approach.

3. Results and Discussion

3.1 Radiation-induced G2 scores and γ-H2AX MFI

Figure 1 shows the radiation-induced G2 (riG2) scores for both controls (a) and patients (b) with each radiation dose. A clear dose dependent variation in G2 score is seen in both healthy controls and patients. A threshold for significance here was considered as $p<0.01$. At each dose, was found that the difference between the data at 0.05Gy and 0.5Gy was statistically significantly different at $p<0.01$ using a two-tailed paired t-test, for both controls and patients. In addition, it was found that the difference in the G2 scores between controls and patients at 0.5Gy, but not 0.05Gy, was statistically significantly different at the same threshold using the same analysis. A more thorough statistical analysis of this data will be forthcoming at the conclusion of follow-up of the prostate cancer patients enrolled in this study.

γ-H2AX MFI for both controls (a) and patients (b) is shown in Figure 2 with respect to radiation dose. Again a dose dependent variation is seen in both healthy controls and patients, although in this case no statistically significant difference was found between measurements at each dose using a two-tailed t-test. Furthermore, the wide variation in γH2AX MFI between individuals is apparent in the extent of the error bars in both patient and healthy control subsets, as has been found previously (31,32).

Although previous studies of single cell lines and extracted PBMCs (20,48,49) have observed a linear dose response in γH2AX foci over the range from 10mGy to 10000mGy, any linearity in dose response in the data in Figure 2 will be obscured by the large inter-individual variation in γH2AX MFI observed using flow cytometry. This could be due to the lack of sensitivity in flow cytometry for measurement of γH2AX foci as compared to assays based on manual or automated foci scoring by microscopy(50). However work by Roch-Lefevre which used fluorescence microscopy with automated scoring of foci provides a reference to which our data can be compared. Interrogation of the coefficient of variation in γH2AX MFI with dose (see table 1) for the control cohort reveals that it is increasing with dose in the control cohort (R^2 (adj.) =0.984), with the level at 0.5Gy similar to that observed previously (coefficient of variation of 0.51) by Roch-Lefevre et al (51).

In respect of the prostate cancer cohort coefficients of variation are very large, and we are not aware of any comparable work here in the literature. In the case of this data, again the variability could be due to a lack of sensitivity of the flow cytometry approach, but equally could be due to the complexity of the cellular response to ionising radiation (52) or indeed the molecular mechanisms by which prostate cancer develops (53). The molecular mechanisms underlying both of these phenomena are not yet fully understood and the genesis of prostate cancer has been attributed on a molecular level to a variety of different genes with different functions. Some of these are involved in DNA damage/repair mechanisms which could affect any experiments examining DNA damage using γH2AX fluorescence.

3.2 Mean Raman spectra of lymphocytes from healthy control and prostate cancer patients

Mean spectra of lymphocytes from both healthy controls and prostate cancer patients are shown in figure 3. A difference spectrum is also provided which shows bands in which intensity differences exist between the two cohorts, including significant differences in the band centred at 783 cm^{-1} (associated with overlapping modes of vibration of both the symmetric stretch of DNA phosphodiester ($–PO_2^-$) and ring breathing modes in uracil, cytosine and thymine) and 1431 cm^{-1} (associated with the overlapping modes of the CH deformation vibration in guanine, adenine and lipids generally)(49,54). In addition an apparent shift in the band at 1001 cm^{-1} in the healthy control to 995 cm^{-1} in the prostate cancer sample was seen, where this band is associated with the phenylalanine ring breathing mode(54). The shading in the bottom panel of figure 3 shows the regions of the difference spectrum which are statistically significant (using a two-tailed t-test at $p<0.001$) between the two samples. It is clear from this data that the lymphocyte fraction of whole blood offers diagnostic capabilities for prostate cancer using Raman spectroscopy. This will be explored in future work.

3.3 Regression of Raman spectra onto radiation-induced G2 scores and γH2AX MFI

The performance characteristics for both PLSR and SVR in predicting riG2 scores and γH2AX MFI is shown in table 2. Both root-mean squared error of calibration (RMSEC) and prediction (RMSEP) are the key outputs and represent the uncertainty in prediction of the target (G2 or γH2AX-MFI) at its training (calibration) and testing (prediction) stage. From table 2 it can be concluded that neither regression algorithm is capable of predicting G2 score with a high degree of confidence. This implies that biochemical signatures within Raman spectra are not predictive of riG2 score and its measure of

radiosensitivity, and that the complex relationship that exists cannot easily be captured by PLSR or SVR.

In contrast, the performance characteristics for prediction of γH2AX MFI from Raman spectra using either PLSR or SVR are shown in table 3. In figures 4 and 5 visualisations of one execution (from the total of 20) of the PLSR and SVR algorithms are also provided. It is clear that both algorithms are capable of predicting the γH2AX MFI target with varying levels of confidence, with the non-linear SVR algorithm the superior overall. This latter result is important and suggests that a non-linear relationship exists between the spectral signatures of DNA double strand break (DSB) damage and repair which exist within the Raman spectra of PBMCs and the measures of the level of DNA DSB represented by γH2AX MFI. In previous work we have demonstrated that a non-linear relationship exists between spectral signatures of keratinocytes and radiation dose in-vitro, where we have characterised this relationship as following the induced repair model for bands associated with both DNA and RNA backbone vibrations (15). While we have found that the intensities of vibrational modes associated with other species do not necessarily follow the induced repair model, it is possible that the non-linearity suggested by the SVR results signifies that more isolated spectral regions that are signatures of DNA DSB's and repair foci do exist, although this would be the subject of a dedicated analysis beyond the scope of the present article. Importantly isolating spectral regions which are associated with DNA DSB repair within the context of the intra-individual variability in γH2AX MFI seen here and elsewhere will potentially require a larger sample set than used in the present work.

Both the PLSR cross-validation results (Figure 4a) and the SVR algorithm parameters (Table 3) demonstrate the complexity of the models that are required to predict γH2AX MFI using either approach. Although model parsimony is desirable in general, the relative agreement between RMSE within the training and testing sets for either PLSR or SVR indicates that the underlying spectral signatures selected by either model are both consistent and robust across the sample set, i.e. that little overfitting by either algorithm occurred. Additionally this level of model complexity has been observed in previous studies of irradiated keratinocytes (14,55).

Finally, the RMSE of prediction (RMSEP) obtained using the SVR algorithm indicates that our method using Raman spectroscopy to predict levels of γH2AX focal damage has an uncertainty in the region of 1.59 MFI (with data normalized to the control, 0Gy sample), which in the context of the data used in the current work where the maximum γH2AX MFI recorded was 32.4, this represents an uncertainty level of approximately 5%, correlating well with previous data on uncertainties in retrospective dosimetry with vibrational spectroscopy (14,55). This gives further encouragement for the use of this method in label free retrospective biodosimetry and other prospective analyses using the PBMC fraction of whole blood.

4. Conclusion

Raman spectroscopy has been suggested as representing a label-free method for the characterisation of biological samples, including in the context of cancer biology and radiobiology. The present work confirms that measurements of the Raman spectra of the PBMC fraction of whole blood may be used to provide quality estimates of the level of DNA damage in humans. Data here also shows

that a link between the Raman spectra of lymphocytes and measures of radiosensitivity from the G2 chromosomal aberration assay is a very complex one which requires further investigation, including in respect of using higher dose points than were used in the present work, and age matched controls. This work therefore confirms that Raman spectroscopy offers a label-free assay for the characterisation of radiobiological response both *in-vitro* and *ex-vivo.*

Acknowledgments

This work was financially supported by the EU FP7 Network of Excellence DoReMi (Grant Number 249689) and Science Foundation Ireland (11/RFP.1/BMT/3317).

Declaration

The authors declare no conflicts of interest.

References

1. Moses RE. Dna damage processing defects and disease. *Annu Rev Genomics Hum Genet* 2001;2:41-68.
2. Thompson LH. Recognition, signaling, and repair of dna double-strand breaks produced by ionizing radiation in mammalian cells: The molecular choreography. *Mutation Research/Reviews in Mutation Research* 2012;751:158--246.
3. Zhou BB, Elledge SJ. The dna damage response: Putting checkpoints in perspective. *Nature* 2000;408:433-9.
4. Jackson SP, Bartek J. The dna-damage response in human biology and disease. *Nature* 2009;461:1071-1078.
5. Coles CE, Moody AM, Wilson CB, et al. Reduction of radiotherapy-induced late complications in early breast cancer: The role of intensity-modulated radiation therapy and partial breast irradiation. *Clinical Oncology* 2005;17:16-24.
6. Coles CE, Moody AM, Wilson CB, et al. Reduction of radiotherapy-induced late complications in early breast cancer: The role of intensity-modulated radiation therapy and partial breast irradiation. *Clinical Oncology* 2005;17:98-110.
7. Herskind C, Talbot CJ, Kerns SL, et al. Radiogenomics: A systems biology approach to understanding genetic risk factors for radiotherapy toxicity? *Cancer Lett* 2016;382:95-109.
8. Kallaway C, Almond LM, Barr H, et al. Advances in the clinical application of raman spectroscopy for cancer diagnostics. *Photodiagnosis Photodyn Ther* 2013;10:207-19.
9. Kong K, Kendall C, Stone N, et al. Raman spectroscopy for medical diagnostics--from in-vitro biofluid assays to in-vivo cancer detection. *Adv Drug Deliv Rev* 2015;89:121-34.
10. Ramos IR, Meade AD, Ibrahim O, et al. Raman spectroscopy for cytopathology of exfoliated cervical cells. *Faraday Discussions* 2016;187:187-198.
11. 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 2016.
12. Horsnell JD, Kendall C, Stone N. Towards the intra-operative use of raman spectroscopy in breast cancer-overcoming the effects of theatre lighting. *Lasers Med Sci* 2016;31:1143-9.
13. Matthews Q, Brolo A, Lum J, et al. Raman spectroscopy of single human tumour cells exposed to ionizing radiation in vitro. *Phys Med Biol* 2011;56:19-38.
14. Meade AD, Clarke C, Byrne HJ, et al. Fourier transform infrared microspectroscopy and multivariate methods for radiobiological dosimetry. *Radiat Res* 2010;173:225-37.
15. Meade AD, Howe O, Unterreiner V, et al. Vibrational spectroscopy in sensing radiobiological effects: Analyses of targeted and non-targeted effects in human keratinocytes. *Faraday Discussions* 2016;187:213-234.
16. Maguire A, Vegacarrascal I, White L, et al. Analyses of ionizing radiation effects in vitro in peripheral blood lymphocytes with raman spectroscopy. *Radiation Research* 2015;183:407-416.
17. Redon CE, Nakamura AJ, Gouliaeva K, et al. The use of gamma-h2ax as a biodosimeter for total-body radiation exposure in non-human primates. *PLoS One* 2010;5:e15544.
18. Bhogal N, Kaspler P, Jalali F, et al. Late residual gamma-h2ax foci in murine skin are dose responsive and predict radiosensitivity in vivo. *Radiat Res* 2010;173:1-9.
19. Greve B, Bolling T, Amler S, et al. Evaluation of different biomarkers to predict individual radiosensitivity in an inter-laboratory comparison--lessons for future studies. *PLoS One* 2012;7:e47185.
20. Horn S, Barnard S, Rothkamm K. Gamma-h2ax-based dose estimation for whole and partial body radiation exposure. *PLoS One* 2011;6:e25113.
21. Horn S, Barnard S, Brady D, et al. Combined analysis of gamma-h2ax/53bp1 foci and caspase activation in lymphocyte subsets detects recent and more remote radiation exposures. *Radiat Res* 2013;180:603-9.

22. Howe O, O'Malley K, Lavin M, et al. Cell death mechanisms associated with g2 radiosensitivity in patients with prostate cancer and benign prostatic hyperplasia. *Radiat Res* 2005;164:627-34.
23. Baria K, Warren C, Roberts SA, et al. Chromosomal radiosensitivity as a marker of predisposition to common cancers? Br j cancer; 2001. pp. 892-6.
24. Papworth R, Slevin N, Roberts SA, et al. Sensitivity to radiation-induced chromosome damage may be a marker of genetic predisposition in young head and neck cancer patients. *Br J Cancer* 2001;84:776-82.
25. Sanford KK, Parshad R, Price FM, et al. Cytogenetic responses to g2 phase x-irradiation of cells from retinoblastoma patients. *Cancer Genet Cytogenet* 1996;88:43-8.
26. Andersson HC, Lewensohn R, Mansson-Brahme E. Chromosomal sensitivity to x-ray irradiation during the g2 phase in lymphocytes of patients with hereditary cutaneous malignant melanoma as compared to healthy controls. *Mutat Res* 1999;425:9-20.
27. Parshad R, Price FM, Bohr VA, et al. Deficient dna repair capacity, a predisposing factor in breast cancer. *Br J Cancer* 1996;74:1-5.
28. Roberts SA, Spreadborough AR, Bulman B, et al. Heritability of cellular radiosensitivity: A marker of low-penetrance predisposition genes in breast cancer? *Am J Hum Genet* 1999;65:784-94.
29. Riches AC, Bryant PE, Steel CM, et al. Chromosomal radiosensitivity in g2-phase lymphocytes identifies breast cancer patients with distinctive tumour characteristics Br j cancer; 2001. pp. 1157-61.
30. Baeyens A, Thierens H, Claes K, et al. Chromosomal radiosensitivity in breast cancer patients with a known or putative genetic predisposition. *Br J Cancer* 2002;87:1379-85.
31. Vral A, Thierens H, Baeyens A, et al. Chromosomal aberrations and in vitro radiosensitivity: Intra-individual versus inter-individual variability. *Toxicol Lett* 2004;149:345-52.
32. Yoon DS, Wersto RP, Zhou W, et al. Variable levels of chromosomal instability and mitotic spindle checkpoint defects in breast cancer. *Am J Pathol* 2002;161:391-7.
33. Ernestos B, Nikolaos P, Koulis G, et al. Increased chromosomal radiosensitivity in women carrying brca1/brca2 mutations assessed with the g2 assay. *Int J Radiat Oncol Biol Phys* 2010;76:1199-205.
34. Scott D, Spreadborough AR, Roberts SA. Less g(2) arrest in irradiated cells of breast cancer patients than in female controls: A contribution to their enhanced chromosomal radiosensitivity? *Int J Radiat Biol* 2003;79:405-11.
35. Parshad R, Sanford KK, Jones GM. Chromatid damage after g2 phase x-irradiation of cells from cancer-prone individuals implicates deficiency in dna repair. *Proc Natl Acad Sci U S A* 1983;80:5612-6.
36. Sanford KK, Parshad R, Gantt R, et al. Factors affecting and significance of g2 chromatin radiosensitivity in predisposition to cancer. *Int J Radiat Biol* 1989;55:963-81.
37. Scott D, Spreadborough AR, Roberts SA. Radiation-induced g2 delay and spontaneous chromosome aberrations in ataxia-telangiectasia homozygotes and heterozygotes. *Int J Radiat Biol* 1994;66:S157-s163.
38. Bryant PE, Gray L, Riches AC, et al. The g2 chromosomal radiosensitivity assay. *Int J Radiat Biol* 2002;78:863-6.
39. Lillicrap SC, B. O, R. WJ, et al. Code of practice for high-energy photon therapy dosimetry based on the npl absorbed dose calibration service. *Physics in Medicine and Biology* 2002;35:1355--1360.
40. Lewis D, Micke A, Yu X, et al. An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan. *Med Phys* 2012;39:6339-50.
41. Micke A, Lewis DF, Yu X. Multichannel film dosimetry with nonuniformity correction. *Med Phys* 2011;38:2523-34.
42. Fiandra C, Ricardi U, Ragona R, et al. Clinical use of ebt model gafchromic film in radiotherapy. *Med Phys* 2006;33:4314-9.
43. Howe OL, Daly PA, Seymour C, et al. Elevated g2 chromosomal radiosensitivity in irish breast cancer patients: A comparison with other studies. *Int J Radiat Biol* 2005;81:373-8.

44. Varmuza K, Filzmoser P. Introduction to multivariate statistical analysis in chemometrics: CRC Press Taylor and Francis Group; 2009.
45. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methodsCambridge: Cambridge University Press; 2000.
46. Smola AJ, Schölkopf B. A tutorial on support vector regression. *Statistics and Computing* 2004;14:199-222.
47. Schölkopf B, Smola AJ, Williamson RC, et al. New support vector algorithms. *Neural Computation* 2000;12:1207-1245.
48. Asaithamby A, Chen DJ, Division of Molecular Radiation Biology DoRO, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA. Cellular responses to dna double-strand breaks after low-dose γ-irradiation. *Nucleic Acids Research* 2017;37:3912-3923.
49. Andrievski A, Wilkins RC. The response of gamma-h2ax in human lymphocytes and lymphocytes subsets measured in whole blood cultures. http://dxdoiorg/101080/09553000902781147 2009.
50. Redon CE, Weyemi U, Parekh PR, et al. Γ-h2ax and other histone post-translational modifications in the clinic. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 2012;1819:743-756.
51. Roch-Lefevre S, Mandina T, Voisin P, et al. Quantification of gamma-h2ax foci in human lymphocytes: A method for biological dosimetry after ionizing radiation exposure. *Radiat Res* 2010;174:185-94.
52. Rosenstein BS, West CM, Bentzen SM, et al. Radiogenomics: Radiobiology enters the era of big data and team science. *Int J Radiat Oncol Biol Phys* 2014;89:709-13.
53. Sreekumar A, Poisson LM, Rajendiran TM, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009;457:910.
54. Meade AD, Lyng FM, Knief P, et al. Growth substrate induced functional changes elucidated by ftir and raman spectroscopy in in-vitro cultured human keratinocytes. *Analytical and Bioanalytical Chemistry* 2007;387:1717-1728.
55. Meade AD, Clarke C, Byrne HJ, et al. Selection of preprocessing methodology for multivariate regression of cellular ftir and raman spectra in radiobiological analyses 2014:254--260.

Figures and Tables

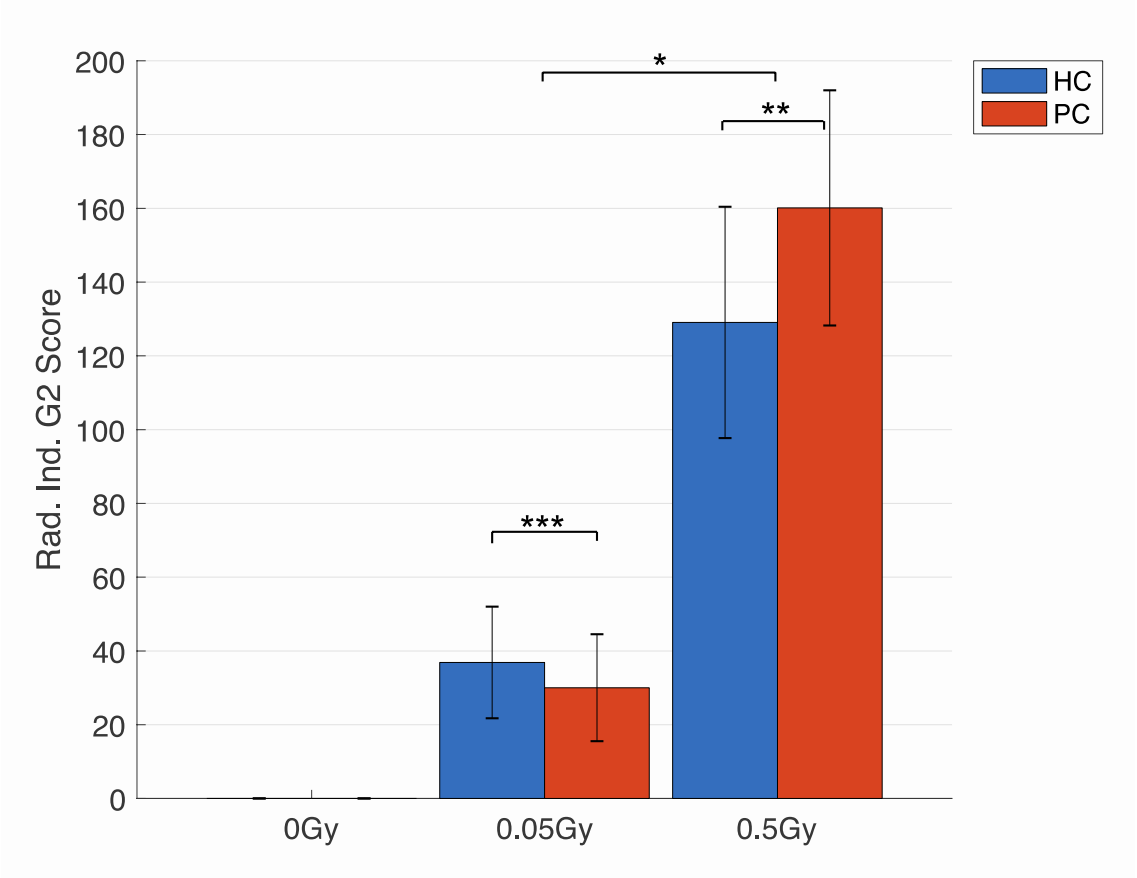

Figure 1: Measured radiation-induced G2 scores with dose in control (HC) and prostate cancer (PC) patients. Error bars depict the standard deviation from the mean at each dose. The radiation-induced G2 scores are normalised by subtraction of the recorded G2 score for any individual at 0Gy from that measured at any dose to correct for spontaneous aberrations. The asterisks indicate the level of significance of the differences recorded from a two-tailed, paired t-test; * = $p<1e-4$; ** = $p<0.005$; *** = $p<0.2$.

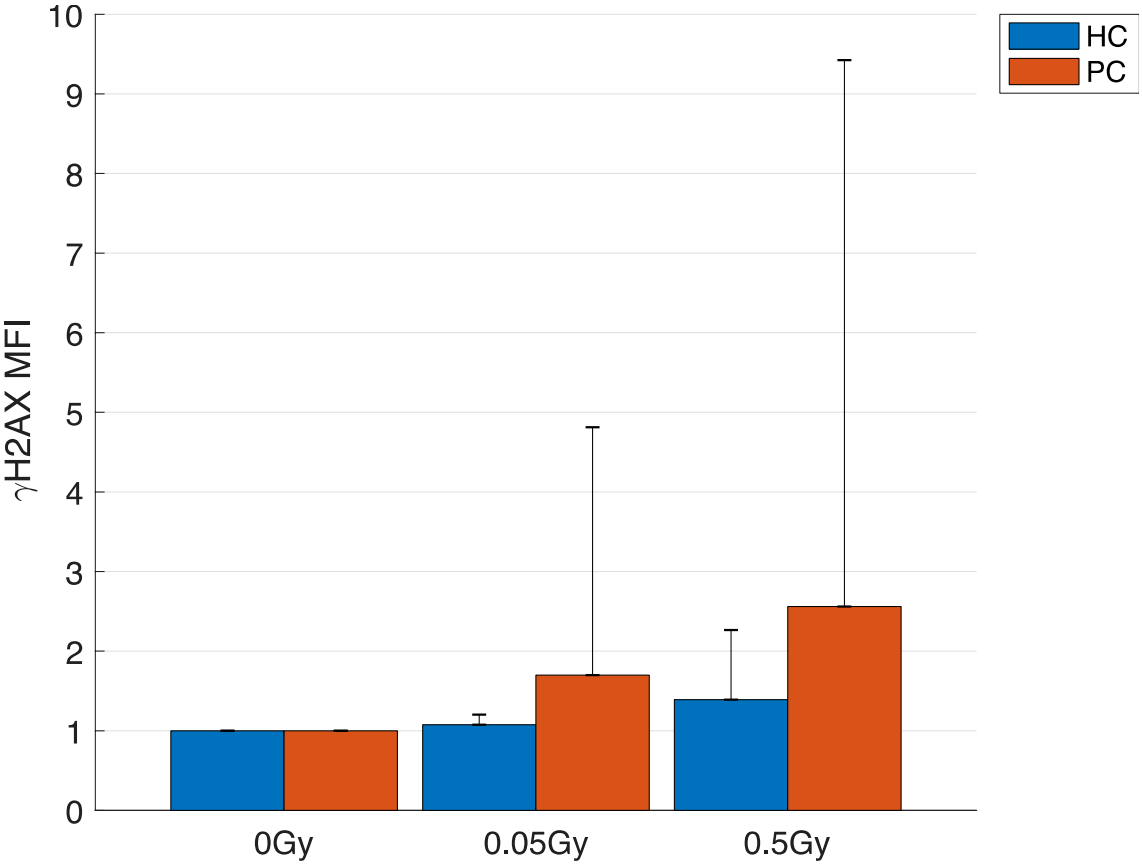

Figure 2. Measured γH2AX mean-fold increase in fluorescence intensity (MFI) (normalized to the 0Gy control) with dose in control (HC) and prostate cancer (PC) patients. Error bars depict the standard deviation from the mean at each dose and are shown in the positive direction only due to their magnitude. No statistically significant differences were observed between γH2AX MFI levels at doses of 0.05Gy and 0.5Gy in either control or prostate cancer patients.

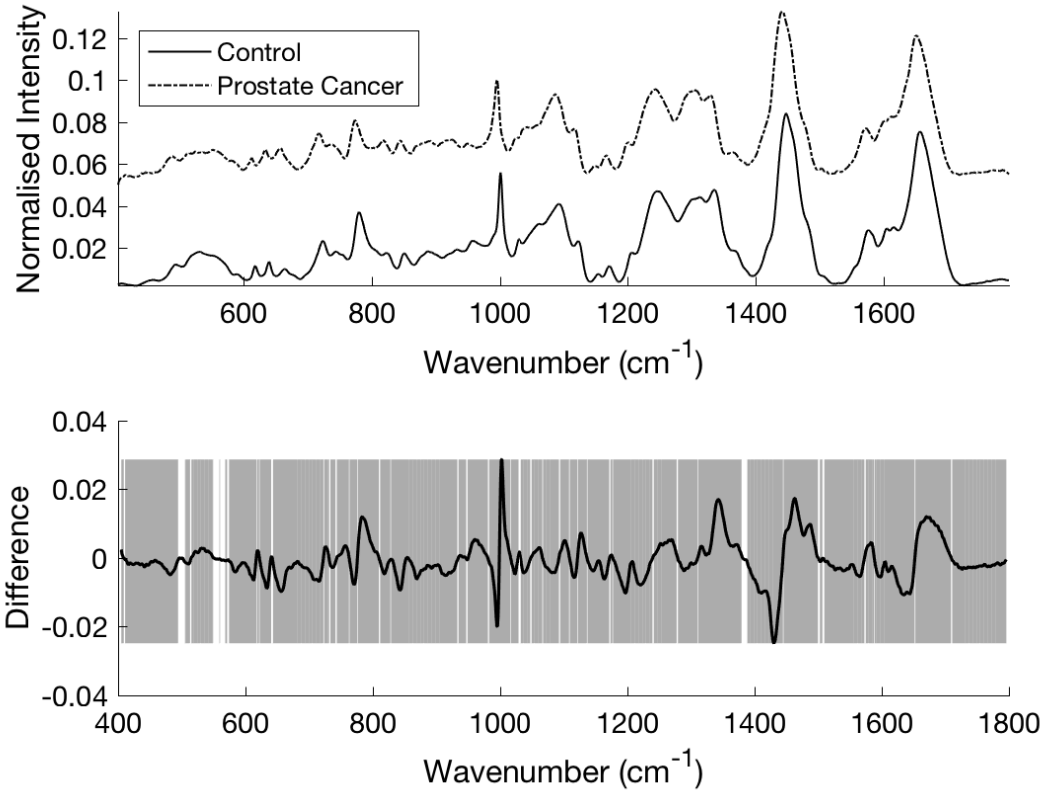

Figure 3. Mean (top) and difference spectra (bottom) for peripheral blood mononuclear cells (PBMCs) for sham-irradiated control and prostate cancer patients. The shaded regions in the bottom panel depict the spectral regions which are significantly different between each sample set using a two-tailed t-test with $p<0.001$.

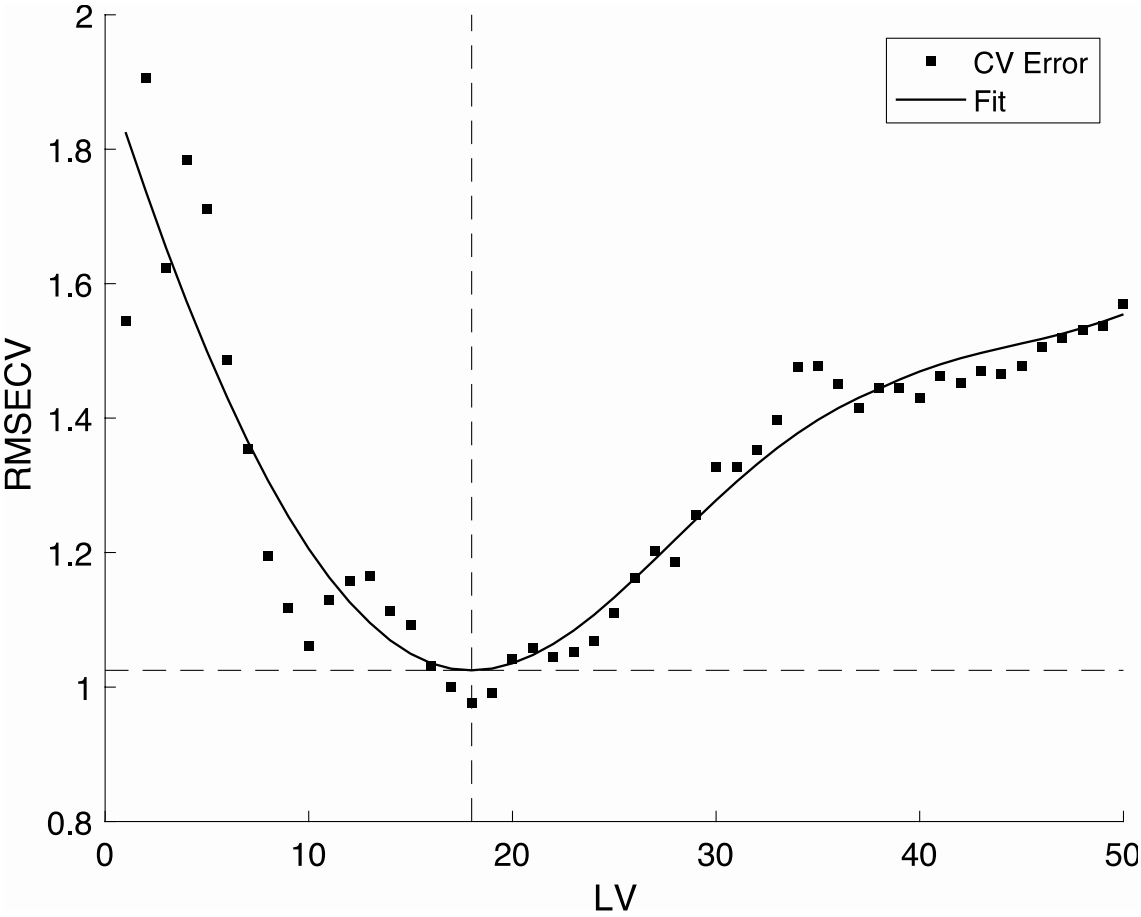

Figure 4a. Sample visualization of cross-validation error (root-mean squared error of cross-validation) of PLSR at training showing optimal model complexity (number of latent variables, LV) at intersection of dashed lines.

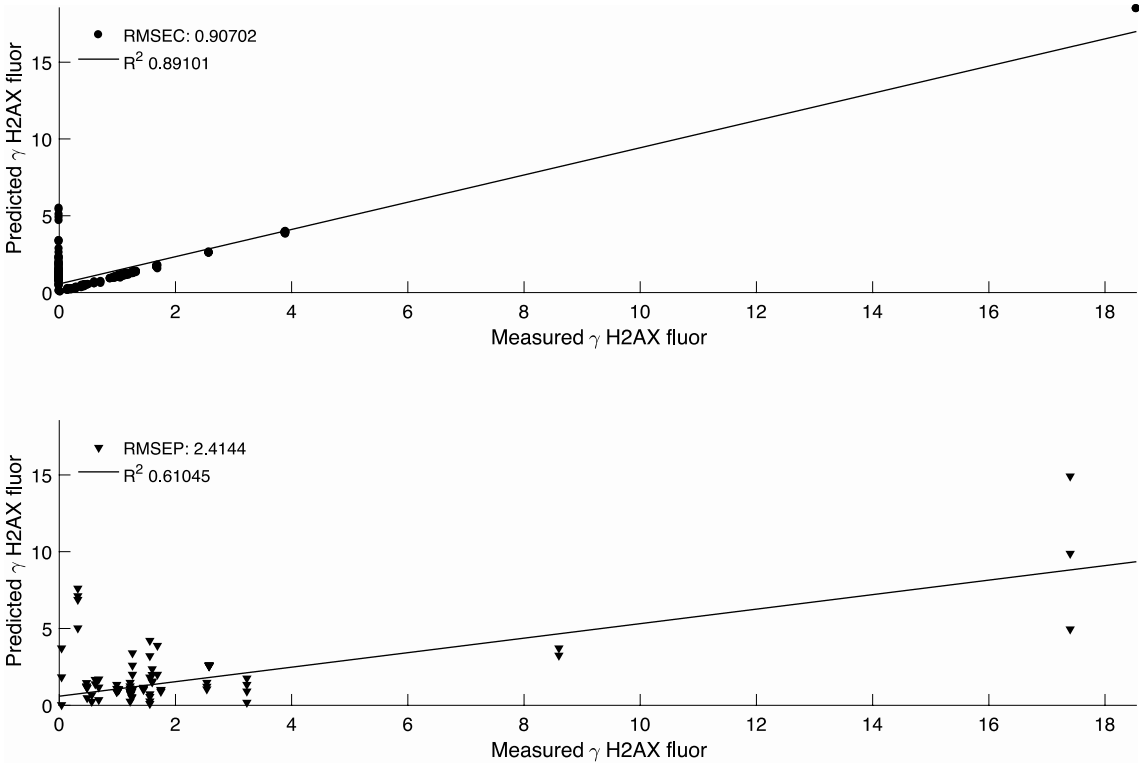

Figure 4b. Sample results of PLSR of PBMC Raman spectra versus γH2AX MFI for (top) training data and (bottom) test data.

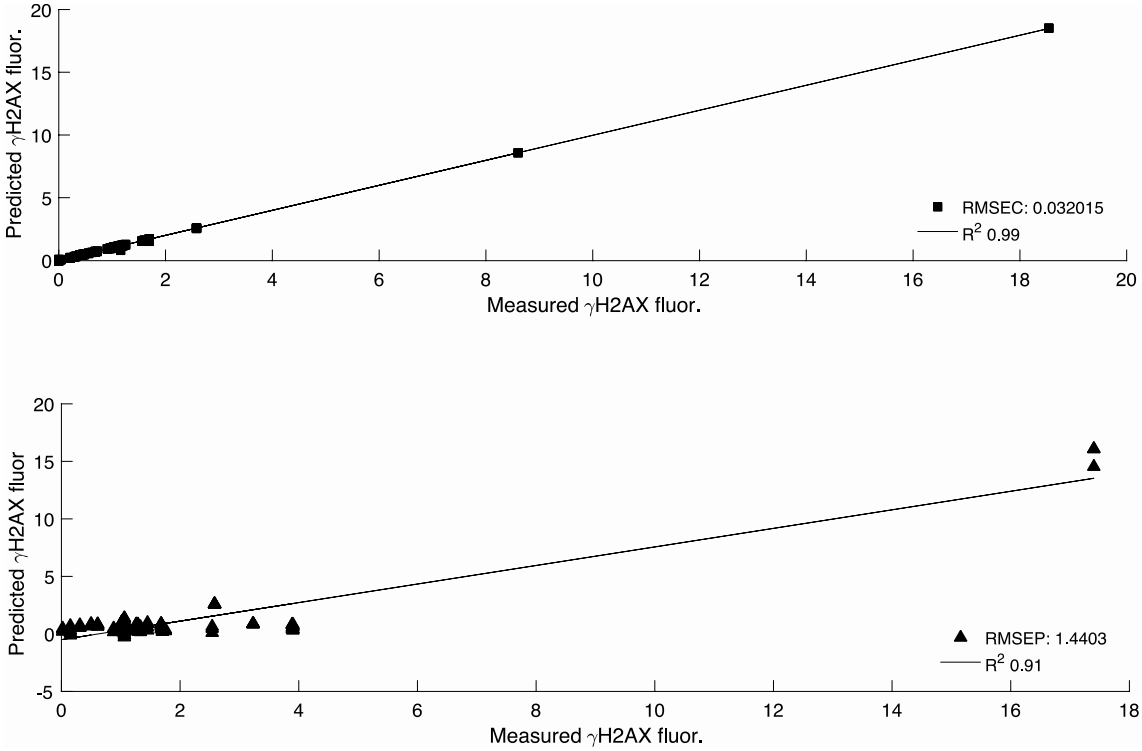

Figure 5. Sample results of SVR of PBMC Raman spectra versus γH2AX MFI for (top) training data and (bottom) test data.

Table 1. Coefficients of variation (CV) in γH2AX MFI with radiation dose for healthy control (HC) and prostate cancer (PC) cohorts.

Dose	**0Gy**	**0.05Gy**	**0.5Gy**
CV HC	0	0.12	0.63
CV PC	0	1.83	2.68

Table 2. Collated performance metrics for both PLSR and SVR of Raman spectra versus radiation-induced G2 score. All are given as the mean (with standard deviation in brackets) for prediction of G2 score using Raman spectra of PMBCs over 20 separate models.

Metric	**PLSR**	**SVR**
RMSEC	55 (5)	26 (4)
RMSEP	118 (18)	86 (14)
R^2 Training	0.41 (0.11)	0.91 (0.03)
R^2 Test	0.1 (0.1)	0.01 (0.2)
Cost	-	50 (0)
Gamma	-	0.28 (0.08)
Nu	-	0.46 (0.16)
NLV	1.5 (1.1)	-

Table 3. Collated performance metrics for both PLSR and SVR of Raman spectra versus γH2AX MFI. All are given as the mean (with standard deviation in brackets) for prediction of γH2AX MFI using Raman spectra of PBMCs over 20 separate models.

Metric	**PLSR**	**SVR**
RMSEC	0.76 (0.17)	0.09 (0.01)
RMSEP	2.67 (0.82)	1.59 (0.17)
R^2 Training	0.94 (0.01)	0.99 (0.002)
R^2 Test	0.37 (0.37)	0.67 (0.31)

Cost	-	40 (17.8)
Gamma	-	0.26 (0.096)
Nu	-	0.24 (0.17)
NLV	30 (11)	-