Check for updates

Detection of radiosensitive subpopulations *ex-vivo* with Raman microspectroscopy

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

[1]School of Physics, Clinical and Optometric Sciences, Technological University Dublin, Dublin, Ireland,
[2]Radiation and Environmental Science Centre (RESC), Physical to Life Sciences Research Hub, FOCAS
Building, Technological University Dublin, Dublin, Ireland, [3]School of Biological, Health and Sport
Sciences, Technological University (TU) Dublin, Dublin, Ireland, [4]Department of Radiation Oncology,
Saint Luke's Radiation Oncology Network (SLRON), St Luke's Hospital, Dublin, Ireland, [5]Cancer Trials
Ireland, Dublin, Ireland

Although significant advances in understanding the molecular drivers of acquired
and inherited radiosensitivity have occurred in recent decades, a single analytical
method which can detect and classify radiosensitivity remains elusive. Raman
microspectroscopy has demonstrated capabilities in the objective classification
of various diseases, and more recently in the detection and modelling of
radiobiological effect. In this study, Raman spectroscopy is presented as a
potential tool for the detection of radiosensitivity subpopulations represented
by four lymphoblastoid cell lines derived from individuals with ataxia
telangiectasia (2 lines), non-Hodgkins lymphoma, and Turner's syndrome.
These are classified with respect to a population with mixed radiosensitivity,
represented by lymphocytes drawn from both healthy controls, and prostate
cancer patients. Raman spectroscopic measurements were made *ex-vivo* after
exposure to X-ray doses of 0 Gy, 50 mGy and 500 mGy, in parallel to radiation-
induced G2 chromosomal radiosensitivity scores, for all samples. Support vector
machine models developed on the basis of the spectral data were capable of
discrimination of radiosensitive populations before and after irradiation, with
superior discrimination when spectra were subjected to a non-linear
dimensionality reduction (UMAP) as opposed to a linear (PCA) approach.
Models developed on spectral data acquired on samples irradiated *in-vitro* with
a dose of 0Gy were found to provide the highest level of performance in
discriminating between classes, with performances of F1 = 0.92 $\pm$ 0.06
achieved on a held-out test set. Overall, this study suggests that Raman
spectroscopy may have potential as a tool for the detection of intrinsic
radiosensitivity using liquid biopsies.

1 Introduction

Cancer incidence is projected to increase globally by 47%, from 18.1 million to 29.5 million annually, over the period from 2020 to 2040 (1). Consequently an associated increase in the numbers of patients requiring radiotherapeutic treatment is expected, if the proportion of patients receiving radiotherapy as part of their care remains constant at ~50% (2). While documented increases in incidence of cancer partially stems from development of technologies for detection and diagnosis of malignancies, the increase in incidence has driven the need for more efficient and effective treatment methodologies. In recent years many efforts have been made in the advancement of personalized treatment in radiotherapy and while some success in this area has been made, mainly due to the advancement of technologies associated with confining dose delivery to malignant regions, less success has been made in advancement in biological assays to assess for inter-individual variation in the response to radiotherapeutic treatment.

The dependence of the individual's response to ionising radiation (IR) exposure on various innate genetic, epigenetic and other characteristics has yet to be fully characterised. Efforts have been made to study individual cases of extreme radiosensitivity in order to elucidate the cellular mechanisms involved in severe cases, in conditions such as Ataxia Telangiectasia, Nijmegen Breakage Syndrome and Fanconi's Anemia, with therapeutic implications for patients undergoing radiotherapy treatment or for individuals subject to occupational and/or accidental irradiation (3). However, it has emerged that a distribution in intrinsic radiosensitivity exists in humans. Coles et al. demonstrated that the level of intrinsic radiosensitivity of the general population follows a Gaussian distribution, encompassing radioresistant, normal and radiosensitive individuals (4, 5). In their description we can understand extreme radiosensitivity in cases such as those associated with Ataxia Telangiectasia, as outlying cases within the distribution of intrinsic radiosensitivity. In addition, the existence of this phenomenon of inter-individual variation highlights the requirement for individual assessment of radiosensitivity to curtail dose delivery to patients undergoing radiotherapeutic treatment. Curtailing dose delivery would result in the reduction of normal tissue complications during radiotherapy, radiation-induced late effects of radiotherapy, and rule out radiotherapy as a treatment modality in cases of extreme radiosensitivity (6).

Ataxia Telangiectasia is a genetic neurodegenerative disorder resulting from a deficiency in the action or regulation of the gene Ataxia Telangiectasia Mutated (ATM), which is involved in the DNA damage response (DDR) pathway. Non-Hodgkin's lymphomas (NHL) are radiosensitive tumours of the blood that have also been associated with deficiency in the repair of DNA damage through the ATM pathways (7, 8). In contrast Turner Syndrome is a cytogenetic condition in females whereby the X chromosome is partly or completely missing. Limited evidence exists that Turner syndrome is associated with increased radiosensitivity (9).

ATM is responsible for the sensing of DNA double strand breaks (DSBs) and plays an important role in pathways associated with the DDR and cell cycle checkpoint arrest (10, 11). ATM is one of the first proteins that localizes to the site of a DSB and phosphorylates the histone H2AX (γ-H2AX) and Ku proteins in response to detection of a DSB (12). These phosphorylated proteins then recruit further repair molecules to the site of the DSB. If unrepaired, DSBs may lead to cell cycle arrest, senescence or more severe outcomes such as genetic instability or programmed cell death (13, 14).

Since ATM phosphorylates H2AX in response to the formation of DSBs, the γ-H2AX assay has been widely used as an assay for the measurement of DSB sensing and repair (15, 16). 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 (17, 18).

An assay that has proven to be more reliable and reproducible in terms of dose response and as a measure of individual radiosensitivity is the *in-vitro* G2 chromosomal radiosensitivity assay (G2 assay) (19). The G2 assay measures, *in-vitro*, the number of chromosomal breaks (either spontaneously induced or induced as a result of exposure to IR) in cells in the G2 phase of the cell cycle before entering mitosis. Other metrics such as the Mitotic Index and Mitotic Inhibition can be calculated from this assay with or without exposure to IR. Mitotic Index is the ratio of the number of cells in metaphase versus the number of cells in interphase. Mitotic Inhibition is a measure of cell cycle checkpoint response to IR and is a measure of the increase or decrease in the number of cells in metaphase following IR when compared to the unirradiated control (20).

A wealth of recent and previous studies using both infrared and Raman spectroscopy have demonstrated that both measurements can identify and interpret radiobiological alterations in exposed cells, tissues and organisms (21–30). Raman spectroscopy uses the inelastic scattering of coherent light from a specimen, where the change in frequency of the light obtained is a fingerprint of the organic biochemistry of the system, and can be used to classify disease states (31), and biochemical processes associated with radiobiological response (32, 33).

In this study, a classification pipeline was developed using Raman spectroscopy of lymphocytes with machine learning approaches to identify radiosensitive cell populations. Here Raman spectra were acquired from lymphocytes drawn from a mixed population, including from the blood of healthy controls and prostate cancer patients, to represent a distribution of radiosensitivity. Spectra were also acquired from lymphoblastoid cell lines established from patients with a range of syndromes conferring innate radiosensitivity (Ataxia Telangiectasia (AT), Turner Syndrome and AT plus NHL (AT-NHL)). Although many spectroscopic studies on the effects of ionising radiation on various cell lines at various radiation doses (34), to the best of our knowledge this study represents the first which has utilized cellular models of radiosensitivity that are well characterised genetically, cytologically and radiobiologically.

Samples were irradiated with doses of 0 Gy, 50 mGy and 500 mGy to test whether innate spectral differences within the resting lymphocytes, or those emerging from irradiation, would provide discrimination of samples based on radiosensitivity. At these dose

levels it is possible to decipher and differentiate on the molecular mechanisms associated with radiation sensitivity across cell and tissue types (35). In this study the radiation-induced G2 score radiosensitivity assay, which utilizes doses up to 500mGy, provided a cytological reference metric of radiosensitivity.

Machine learning classification models employing these spectra were developed using a support vector machine (SVM) algorithm to which spectral data was input after decomposition using principal component analysis (PCA) or universal manifold approximation and projection (UMAP). Our results indicate that incorporation of a UMAP decomposition step within the pipeline provides superior model performance to that achieved using PCA, suggesting a strategy for spectral pre-treatment within models for identification of radiosensitive individuals in practice.

2 Materials and methods

Culture of lymphoblastoid cell lines

All cell lines used in this study were lymphoblastoid cell lines (LCL's) with EBV-immortalised B-lymphocytes and they were generated from whole blood samples derived from individuals. Two were obtained as a gift from the Institut Curie, Paris, France, and found to be consistent with a diagnosis of non-Hodgkins lymphoma and Turner's syndrome (henceforth collectively labelled as NHL-T cell lines; first described by Angéle et al. (36)). Another two cell lines were obtained as gifts from the University of Birmingham and were derived from two distinct AT-patients (henceforth collectively labelled as AT lines). These lines have also been described previously (37).

LCL's are cells which grow in suspension and do not attach to surfaces within their environment. They were maintained in standard RPMI medium (Sigma) supplemented with 10% Foetal bovine serum (FBS), L-Glutamine and incubated at 37°C with 5% CO_2. Each cell line was kept at a density of 1×10^5/ml and split by 1:5 dilutions every 24-48 hrs. Prior to irradiation, 5 ml of cells at a density of 1×10^6 cells/ml were seeded per flask, per cell line and per dose.

Isolation of lymphocytes, blood culture and irradiation

This study received ethical approval from the Technological University Dublin Research Ethics Committee (REC number 15-32). Blood was drawn from a total of 23 healthy controls and 19 prostate cancer patients (prior to treatment) as described previously (28, 38). Prostate cancer patients were recruited under a Cancer Trials Ireland trial, CTRIAL-IE (ICORG) 08-17 (NCT00951535). Peripheral blood mononuclear cells were then isolated as described in our previous work (28) through establishment of a blood culture for each blood draw, and isolation of lymphocytes for irradiation through plastic adherence over a 72-hour culture period.

Sample irradiation

In-vitro cultured lymphocytes (for Raman spectroscopy) and blood cultures (for the G2 assay) were both irradiated with an X-ray linear accelerator (LINAC) at SLRON St. Luke's Hospital, Dublin, as described previously (28). Cells were either sham irradiated or irradiated with a 50 mGy or 500 mGy dose. Samples were then incubated for a 1-hour period at 37 °C before sample preparation for spectroscopy.

Radiation-induced G2 assay

The G2 assay was performed as described previously (28), using whole blood cultures. Briefly, cells were irradiated with an X-ray LINAC at a dose of 500 mGy, were arrested in metaphase via incubation with colcemid, and were subsequently fixed, lysed and stained using a 3% Giemsa solution. The number of chromatid aberrations was then recorded microscopically for 50 cells per slide. For each donor, a radiation-induced G2 (RIG2) score was recorded as the difference in the number of aberrations seen in the control (0 Gy) sample versus that in the irradiated sample. Examplar images of the chromosomal aberrations observed during this analysis have been published previously (in White at al, and Bryant et al. (37, 39)).

Raman slide preparation

Following IR exposure, lymphocytes in suspension were first centrifuged at 250 g for 5 minutes. Supernatant was removed and the cells were fixed using 200 μls of 2% paraformaldehyde in phosphate-buffered saline. From the suspension, 40 μl was drop cast onto calcium fluoride (CaF$_2$) slides. Paraformaldehyde was removed using disposable graduated pipettes and the slides for Raman spectroscopy were then rinsed in deionised H$_2$O for 5 minutes. Washing was performed three times and then the slides were allowed to dry for Raman spectroscopic measurements.

Raman spectroscopy and post-processing

Raman spectroscopy was performed using a Horiba Jobin Yvon Labram HR800 UV spectrometer. Spectra were acquired with a 660 nm solid-state diode laser delivering 100 mW output power, with a 20 second integration time and averaged across three integrations per spectrum. Spectra were recorded using a diffraction grating ruled with 300 lines/mm giving a spectral resolution of ~2.1 cm^{-1}. The confocal hole was set to 150 μm with the grating centered at 1350 cm^{-1}. A total of 30 to 50 spectra for each donor, cell line and dose were recorded from the nuclear portion of the cell using a raster scanning methodology described elsewhere (28, 38, 40). The cells were ~8-12 μm in size and each spectrum was recorded from individual cells with a 4x4 μm raster scan of the centre of the cell including both the nucleus and cytoplasm of the cell. All spectra were recorded within two weeks of slide preparation. Spectra were
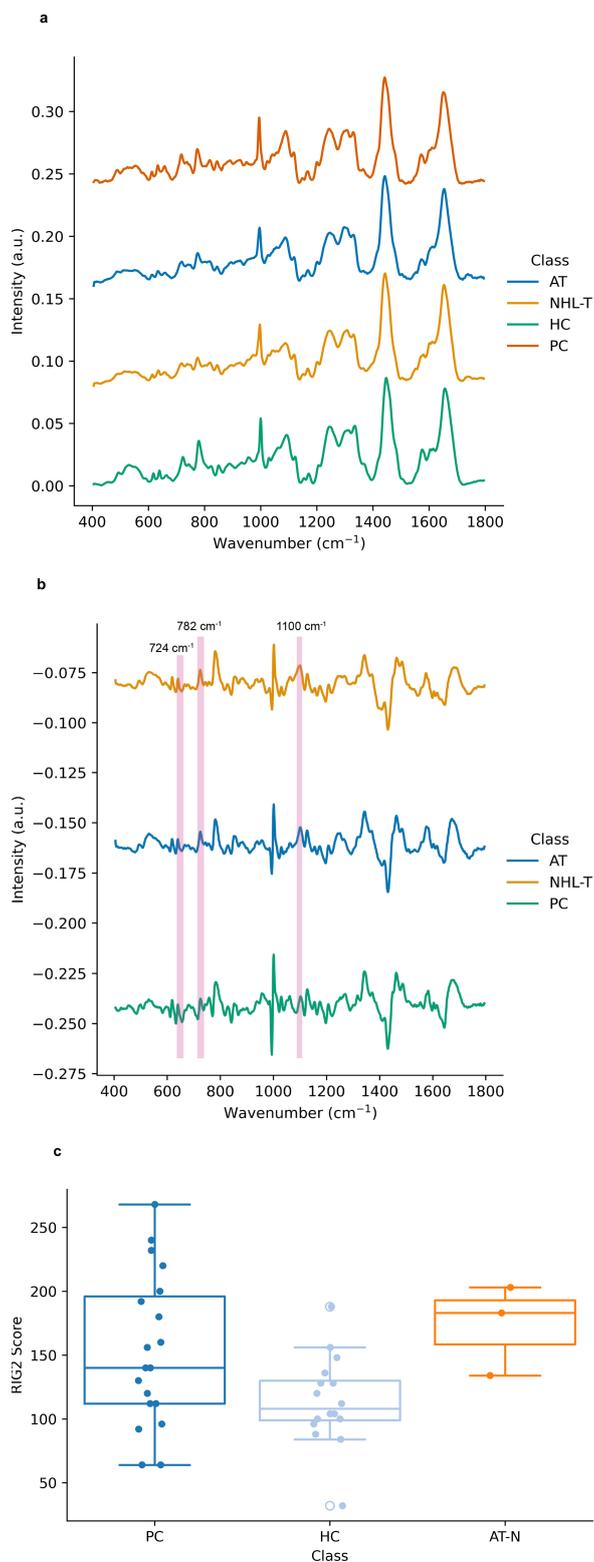
FIGURE 1
(A) Mean Raman spectra of lymphocytes from control samples (0 Gy), including samples drawn from healthy controls (HC), prostate cancer patients (PC), AT-deficient lymphoblastoid cell lines (AT) and two additional immortalized lymphoblastoid cell lines (NHL-T) as described in the main text. Spectra are vertically offset for visual clarity. **(B)** Mean Raman difference spectra for prostate cancer patients (PC), AT-deficient lymphoblastoid cell lines (AT) and reference immortalized lymphoblastoid cell lines (NHL-T) as described in the text. Spectra are offset for visual clarity. **(C)** Distribution of radiation induced G2 scores for each cell subclass. Here the data for the AT and NHL-T cell lines are grouped as 'AT-N'.

downsampled using intersample averaging to produce approximately 8-15 high SNR spectra per donor, cell line and treatment condition, resulting in a total of 1279 spectra for analysis (28).

All post processing was performed using Matlab version 7.9.0 (R2009b; Mathworks, USA) using the PLS-Toolbox version 6.51 (Eigenvector Research Inc.) and in-house algorithms. Wavenumber calibration, intensity calibration, baseline correction, smoothing and removal of substrate contribution were performed as outlined in our previous studies (28, 38). All spectra were subsequently vector normalized before analysis and were downsampled to provide a reduced dataset, improving the overall signal-to-noise ratio in each spectrum. In total 1279 spectra were used for model development.

Machine learning

All analysis and machine learning was conducted in Python v.3.10.9 using sci-kit-learn v.1.4.1.

The dimensionality of the spectral data was reduced before the development of classification models using two approaches, PCA and UMAP. PCA is well established in the field, allowing for the reduction of a spectral dataset from a large number of covariant wavenumbers to a, typically, small number of orthogonal basis variables which describe the underlying variance in the data (38, 41–43). While PCA does allow subsequent interpretation of the origin of spectral discrimination between classes, often this approach is insufficient to reduce the data to a manageable component subset (29, 44, 45), with the result that models can be

overly complex. Recently it has been demonstrated that advanced manifold approaches, including UMAP (46, 47) can be employed with spectral data for variable reduction (44, 48, 49), which, while not in themselves providing insights into the spectral origin of clustering, are effective in reducing spectral data to a small number of components for cluster visualization and model development.

Classification models were developed using a SVM algorithm with either PCA or UMAP as a dimensionality reduction step. The dimensionality reduction step involved the development of a reduction model on the training set which was then applied to both the training and testing data, in order to prevent model leakage via conditioning.

Hyperparameter optimization was performed using a brute-force grid-search cross-validation approach (with 3 folds) for both the SVM algorithm and the dimensionality reduction approaches as follows:

- SVM hyperparameters: C (regularization parameter) – values 0.1,0.5 or 1; kernel – linear, rbf or sigmoid; gamma (penalty parameter) – scale or auto;
- PCA: number of components varied between 1 and 50;
- UMAP: number of components varied between 1 and 4; minimum distance value – 0.1, 0.5 or 0.7; number of neighbours – 30, 100, 170, 230; distance metric – minkowski, Manhattan, Canberra, cosine, or correlation.

All models were executed for 30 separate classification epochs, with data randomized between training (80% of data) and test sets (20% of data) on each pass. Samples were randomized in such a way to ensure model leakage did not occur and that spectra of cells from

FIGURE 2
(A) Principal component analysis score plot for Raman spectra of the control sample (0 Gy) using a model including the first 5 principal components. This model describes a total of 75% of the spectral variance in the data. **(B)** Score plot depicting the typical distribution of UMAP scores obtained from an embedding model with two components.
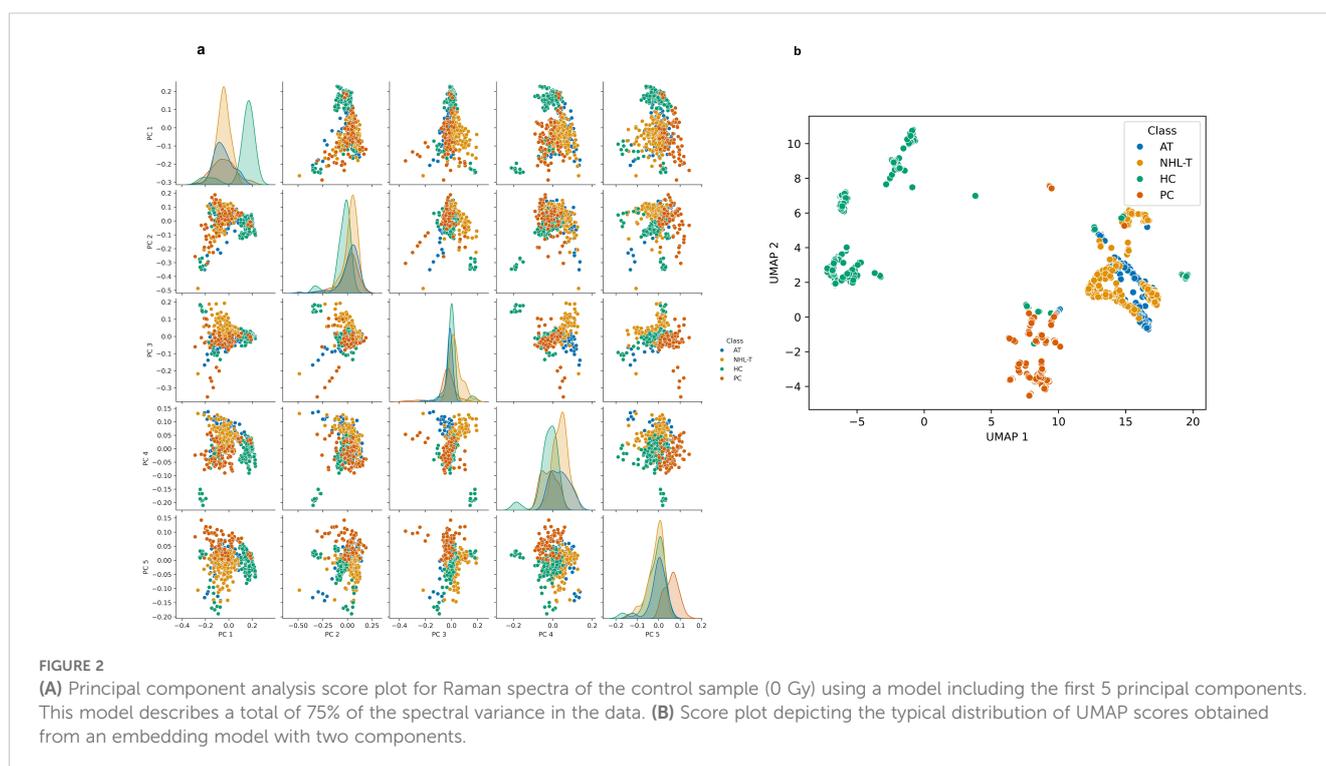
a given line or donor did not appear in both the training and test sets. In the case of the cell lines on each pass it was ensured that spectra from one AT line, and from one or other of the NHL-T lines, did appear in the training set with the other appearing in the test set.

3 Results and discussion

In Figure 1A the mean spectra of each of the cell subclasses are depicted. In Figure 1B the mean difference spectra of each sample class are calculated, where the mean spectrum of lymphocytes
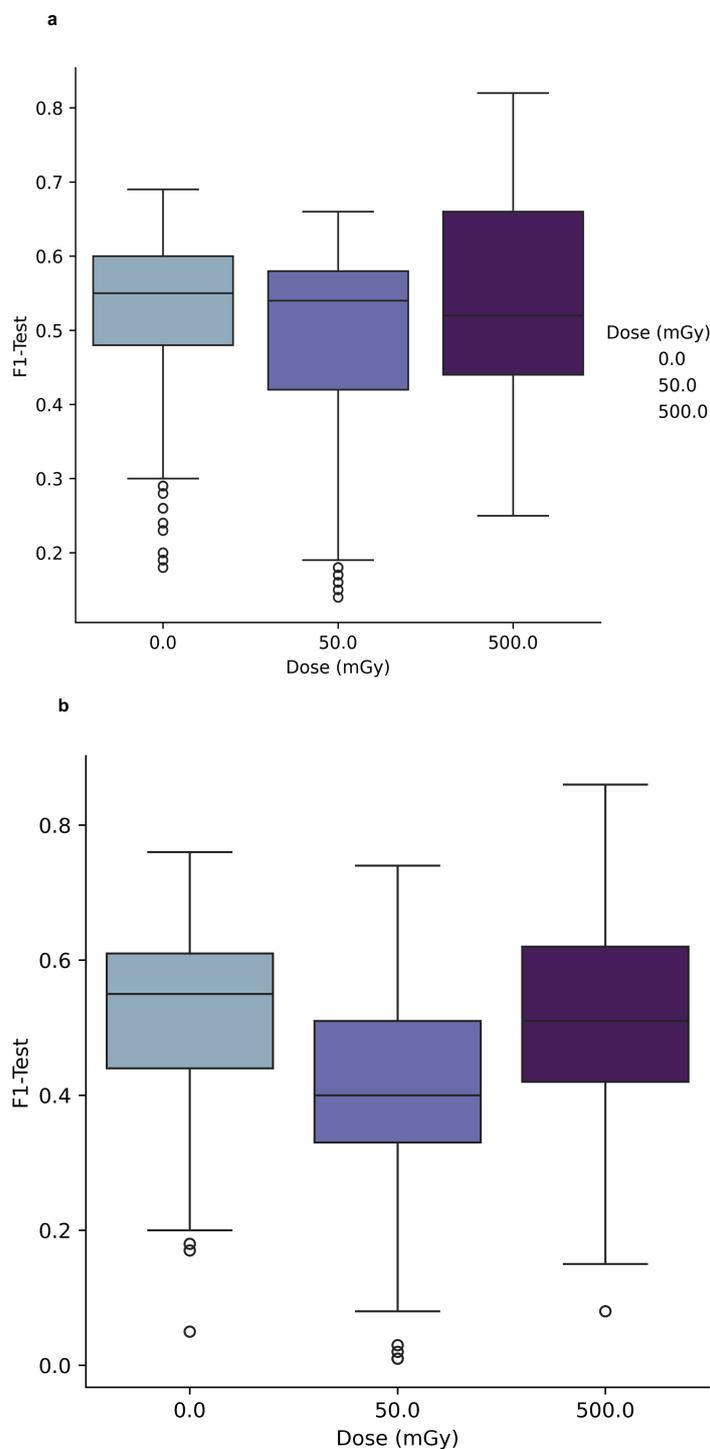
FIGURE 3
(A) Performance of PCA-SVM models as a function of PC and *in-vitro* radiation dose for a 4-class classification model. **(B)** Performance of UMAP-SVM models as a function of *in-vitro* radiation dose for a 4-class classification model.

derived from healthy controls are used as a reference. Here there are significant differences in spectral intensity at 724 cm^{-1} (adenine ring breathing mode), 782 cm^{-1} (uracil, cytosine and tyrosine ring breathing modes) and 1100 cm^{-1} (stretching vibrations in $-PO_2^-$), together with a spectral shift of ~3 cm^{-1} between the normal position of the phenylalanine ring breathing mode (~1004 cm^{-1}) and that seen in the cancer and radiosensitive cell lines. Taken together these features may have their origins in the differences in DNA repair propensities in each of these subclasses relative to the healthy controls and may therefore be viewed as a spectral biomarker of abnormal repair. As transfection by EBV was used to immortalise these cell lines, this may also contribute to the spectral differences observed in the lymphoblastoid cell lines here.

Figure 1C depicts the RIG2 scores measured in these samples (originally partially produced in our previous work (28)), together with RIG2 scores measured for both the healthy control population and the AT and N cell lines (which are grouped together for the purposes of visualisation due to there being RIG2 score data from only 3 cell lines available). While this data exemplifies the range of radiosensitivities which are yielded from cytogenetic approaches in cancer patients, radiosensitive subpopulations and healthy controls, the overlap in the distributions also demonstrates the challenges in using RIG2 scores for the purposes of precision identification of radiosensitive sub-populations.

Figure 2A depicts the PCA scores plot for spectral data at a dose of 0Gy. This PCA model was found to account for 75% of the variance in the spectral data (for reference, a model including the first 10 components was found to describe 86% of the variance in the spectral data). This demonstrates that PCA models account for a substantial proportion of the spectral variance, with a level of clustering seen between scores using many PCs between PC1 and PC5 (Supplementary Figures S1, S2 depict the PCA score plots observed at doses of 50mGy and 500mGy, respectively). However, as the boundary between classes is highly non-linear, this suggests the requirement for deployment of non-linear machine learning approaches for classification here.

Figure 2B then depicts a typical plot of UMAP scores obtained from an embedding utilizing two components, again for spectral data observed at a dose of 0Gy. UMAP score plots, unlike PCA score plots, are highly dependent upon the embedding

hyperparameters, such that the score plot will differ substantially for other embeddings. As a non-linear embedding approach UMAP has been demonstrated to produce robust embeddings which have been shown to preserve the global structure and continuity of datasets (47, 50). Here it can be seen that this non-linear dimensionality reduction approach can produce embeddings which separate spectral classes such that modelling with a non-linear classification approach is likely to produce higher classification performances.

Figures 3A and B, the classification performance of the PCA-SVM and UMAP-SVM models are depicted as a function of *in-vitro* radiation dose for a 4-class classification approach (classes – healthy control, prostate cancer, AT-LCL and N-LCL). Modelling performance was expressed as F1-score on the held-out test set, and provided as the macro-average across each class.

For PCA-SVM and UMAP-SVM models, macro F1-score performances were high ($\mu = 0.96$ ($\sigma \pm 0.07$) and $\mu = 0.95$ ($\sigma \pm 0.02$), respectively) though these reduced significantly when the tuned models were presented data from the held-out test set, with mean PCA-SVM modelling performance for the test set at 0Gy at F1 = 0.54 ($\pm$ 0.08) and that for the UMAP-SVM models averaging at F1 = 0.52 ($\pm$ 0.1) across all hyperparameters. It can be seen from Figures 3A and B that little variation in modelling performance with either dose or hyperparameter for either model type was observed.

To further interrogate the potential to discriminate radiosensitive subpopulations, the AT and N LCL cell lines were grouped together as a unified class and models were then trained separately for discrimination of each class. Results of this analysis are depicted in Figures 4 and 5.

In Figure 4 the test results of a 3-class PCA-SVM classification approach are shown by variation in *in-vitro* radiation dose and number of PCs retained in the model. All models optimize with in the region of 6-7 PCs, with spectral variance increasing with dose, and optimal model performance (mean F1 = 0.89 ($\sigma \pm 0.07$) observed for models trained on spectra at 0Gy, with model complexity set to 6 PCs.

Likewise in Figures 5A, B, the UMAP-SVM model performances are provided with variation in hyperparameter. Modelling performances are again maximized for spectral data at a 0Gy dose, with performance falling with dose. Additionally UMAP-SVM model performances appear to be slightly more robust and parsimonious
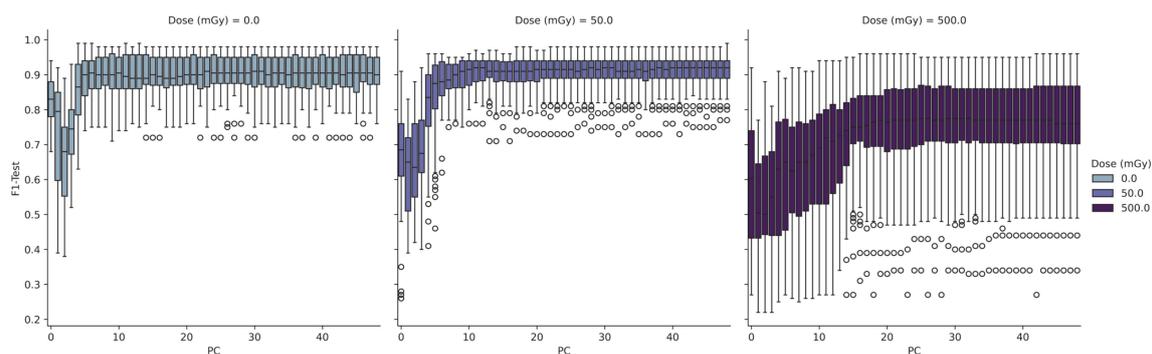
FIGURE 4
Performance of PCA-SVM models as a function of *in-vitro* radiation dose and model complexity (number of PCs).
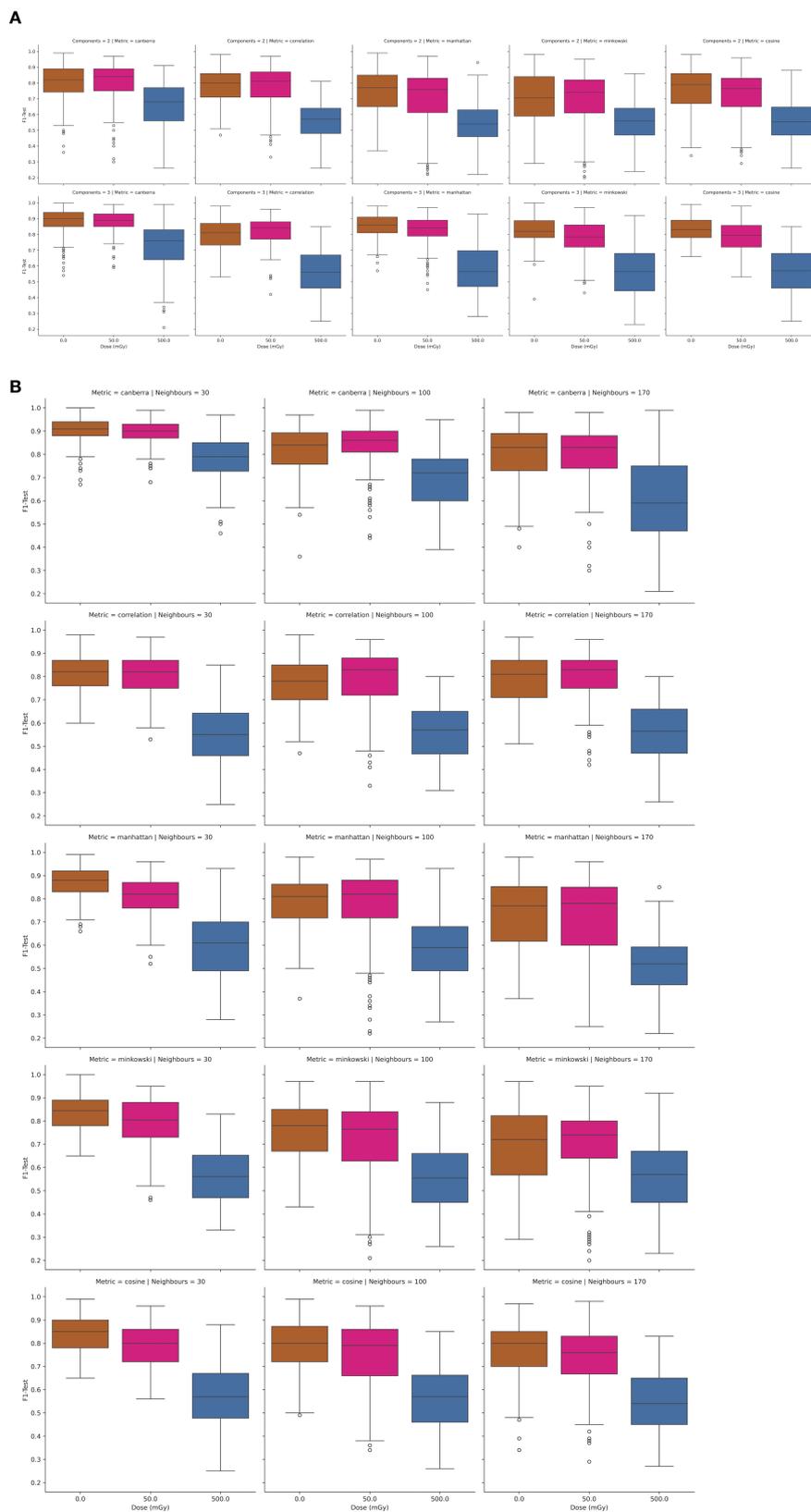
FIGURE 5
(A, B) Performance of UMAP-SVM models as a function of in-vitro radiation dose, UMAP model complexity (number of components, number of neighbours), and distance metric.

when developed using the *canberra* metric (with number of neighbours set to 30, components to 3) with a mean prediction performance on the held-out test set of F1 = 0.92 (σ ± 0.06). A two-tailed t-test of the F1-score performances returned by the models at testing revealed that they were statistically significantly different (p<0.025), suggesting that the UMAP embedding of spectral data provides superior classification performance for this use case.

Within the context of this study, which uses liquid biopsies from healthy individuals, those with a confirmed diagnosis of cancer, and those from individuals with a confirmed genetic diagnosis of conditions conferring radiosensitivity, this high classification performance is encouraging.

This result suggests that spectroscopic assays for radiosensitivity may potentially have an important role to play within a pipeline of assays for identification of individuals at risk of adverse effects from exposure to ionizing radiation, either in a clinical or non-clinical context. Within a clinical context, these findings point to the potential of this spectroscopic assay within a pipeline for a-priori segregation of those individuals likely to possess a radiosensitive phenotype, and who therefore require careful monitoring for adverse responses during radiotherapeutic treatment or may be candidates for alternative therapeutic strategies. Importantly, the success of this spectroscopic assay does rely upon careful optimisation of sample preparation for optimal classification performance (34, 43). Equally, establishing coherent, robust linkages between spectral observations and the underlying molecular events which give rise to the spectral phenotype is crucial to provide confidence to biomedical scientists and clinicians for interpretation of the spectroscopic measurement.

While this experimental finding does require validation in a larger cohort, it is potentially important as, despite a significant body of research in this area, a single assay for radiosensitivity has yet to emerge (35). It is therefore likely that any future deployment of this technology in radiation science will involve its use alongside measurements for other characteristics of radiation response.

4 Conclusion

In this article, AT lymphoblastoid cell lines established from blood samples taken from two different AT patients, a further one with NHL, and a separate lymphoblastoid cell line from an individual with Turner's syndrome were used as cellular models of radiosensitivity, together with spectra of lymphocytes drawn from healthy controls and prostate cancer patients. Parallel reference measurements of cellular radiosensitivity were recorded as ground truth metrics, establishing the difficulties in using, for example, cytogenetic metrics such as RIG2 score as an objective measure of radiosensitivity in mixed human populations.

The study has demonstrated the capability of Raman spectroscopy as a tool for the assessment of individual radiosensitivity. Differences in spectral biochemical signatures o cells have been shown to allow for detection and discrimination of cells based on intrinsic factors relating to radiosensitivity. Of course, this work uses only a few biological examples of radiosensitivity as models, and a caveat must be applied to the fact that lymphoblastoid cell lines, rather than lymphocytes drawn

from patients, have been used as radiosensitive samples here. Overall however, the study demonstrates clearly the capability of the classification pipeline for the identification of radiosensitive cell subpopulations within a wider heterogenous population.

Using a UMAP-SVM model an F1-score of 0.92 ± 0.06 was observed from classification models using Raman spectra acquired on a 0 Gy (control) sample. Therefore, the incorporation of UMAP spectral decomposition within the classification pipeline, and careful hyperparameter tuning, appears to be a key enabling step providing decomposition of spectral data towards classification by machine learning.

This study has highlighted the potential of a measurement of the Raman spectra of a pre-treatment blood biopsy for identification of radiosensitive subpopulations. However, it is likely that further model training data, together with models which utilize spectral data together with other biological metrics, will be required for high-precision isolation of radiosensitive subpopulations in practice.

Data availability statement

Spectral data acquired from the lymphoblastoid cell lines used in this study are available on Zenodo at the following link: https://doi.org/10.5281/zenodo.14888782. The remaining spectral data may be shared by the corresponding author upon reasonable request.

Ethics statement

The studies involving humans were approved by Technological University Dublin Research Ethics Committee (REC number 15-32). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ADM: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. AM: Data curation, Investigation, Writing – original draft. JB: Data curation, Investigation, Writing – original draft. DC: Data curation, Investigation, Writing – original draft. DM: Data curation, Investigation, Writing – original draft. LW: Data curation, Investigation, Writing – original draft. JA: Funding acquisition, Project administration, Resources, Writing – original draft. MD: Data curation, Writing – original draft. EN: Project administration, Writing – original draft. SB: Project administration, Writing – original draft. MF: Project administration, Writing – original draft. AS: Project administration, Writing – original draft. OH: Funding acquisition, Resources, Supervision, Writing – original draft. FL: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2025.1470431/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2021) 71:209–49. doi: 10.3322/caac.21660

2. Baumann M, Ebert N, Kurth I, Bacchus C, Overgaard J. What will radiation oncology look like in 2050? A look at a changing professional landscape in Europe and beyond. *Mol Oncol*. (2020) 14:1577–85. doi: 10.1002/1878-0261.12731

3. Moses RE. DNA damage processing defects and disease. *Annu Rev Genomics Hum Genet*. (2001) 2:41–68. doi: 10.1146/annurev.genom.2.1.41

4. Coles CE, Moody AM, Wilson CB, Burnet NG. Reduction of radiotherapy-induced late complications in early breast cancer: the role of intensity-modulated radiation therapy and partial breast irradiation. *Clin Oncol*. (2005) 17:16–24. doi: 10.1016/j.clon.2004.08.008

5. Coles CE, Moody AM, Wilson CB, Burnet NG. Reduction of radiotherapy-induced late complications in early breast cancer: the role of intensity-modulated radiation therapy and partial breast irradiation. Part II–Radiotherapy strategies to reduce radiation-induced late effects. *Clin Oncol R Coll Radiol G B*. (2005) 17:98–110. doi: 10.1016/j.clon.2004.08.010

6. Barnett GC, Deborah T, Laura F, Sarah K, Chris T, Er M, et al. A genome wide association study (GWAS) providing evidence of an association between common genetic variants and late radiotherapy toxicity. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. (2014) 111:178–85. doi: 10.1016/j.radonc.2014.02.012

7. Rendleman J, Yevgeniy A, Boris R, Christina A, Przybylo AJ, Dutra-Clarke A, et al. Genetic variation in DNA repair pathways and risk of non-Hodgkin's lymphoma. *PloS One*. (2014) 9:e101685. doi: 10.1371/journal.pone.0101685

8. Justice T. Safety and efficacy of external beam radiation therapy for non-Hodgkin lymphoma in patients with prior 90Y-ibritumomab tiuxetan radioimmunotherapy. *Cancer*. (2006) 107:433–8. J M, G W, T W. doi: 10.1002/cncr.v107:2

9. Heras JG, Coco R. Chromosomal sensitivity to X-rays in lymphocytes from patients with Turner syndrome. *Mutat Res*. (1986) 160:33–8. doi: 10.1016/S0027-5107(96)90006-8

10. Burma S. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J Biol Chem*. (2001) 276:42462–7. P CB, M M, a K, J CD. doi: 10.1074/jbc.C100466200

11. Iliakis G, Ya W, Jun G, Huichen W. DNA damage checkpoint control in cells exposed to ionizing radiation. *Oncogene*. (2003) 22:5834–47. doi: 10.1038/sj.onc.1206682

12. Hoeijmakers JHJ. Genome maintenance mechanisms for preventing cancer. *Nature* (2001) 411:366–74. doi: 10.1038/35077232

13. d'Adda di Fagagna F. Living on a break: cellular senescence as a DNA-damage response. *Nat Rev Cancer*. (2008) 8:512–22. doi: 10.1038/nrc2440

14. Jackson SP. Sensing and repairing DNA double-strand breaks. *Carcinogenesis*. (2002) 23:687–96. doi: 10.1093/carcin/23.5.687

15. Redon CE, Nakamura AJ, Gouliaeva K, Rahman A, Blakely WF, Bonner WM. The use of gamma-H2AX as a biodosimeter for total-body radiation exposure in non-human primates. *PloS One*. (2010) 5:e15544. doi: 10.1371/journal.pone.0015544

16. Bhogal N, Kaspler P, Jalali F, Hyrien O, Chen R, Hill RP, et al. Late residual gamma-H2AX foci in murine skin are dose responsive and predict radiosensitivity *in vivo*. *Radiat Res*. (2010) 173:1–9. doi: 10.1667/RR1851.1

17. Greve B, Bolling T, Amler S, Rossler U, Gomolka M, Mayer C, et al. Evaluation of different biomarkers to predict individual radiosensitivity in an inter-laboratory comparison–lessons for future studies. *PloS One*. (2012) 7:e47185. doi: 10.1371/journal.pone.0047185

18. Horn S, Barnard S, Brady D, Prise KM, Rothkamm K. 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. doi: 10.1667/RR13342.1

19. Howe OL, Daly PA, Seymour C, Ormiston W, Nolan C, Mothersill C. Elevated G2 chromosomal radiosensitivity in Irish breast cancer patients: a comparison with other studies. *Int J Radiat Biol*. (2005) 81:373–8. doi: 10.1080/09553000500147642

20. Cadwell KK. G2 checkpoint control and G2 chromosomal radiosensitivity in cancer survivors and their families. *Mutagenesis*. (2011) 26:291–4. B CG, J TE, F WJ, D BJ. doi: 10.1093/mutage/geq087

21. Gault N, Lj L. Infrared microspectroscopic characteristics of radiation-induced apoptosis in human lymphocytes. *Radiat Res*. (2003) 160:238–50. doi: 10.1667/RR3020.1

22. Meade AD, Clarke C, Byrne HJ, Lyng FM. Fourier transform infrared microspectroscopy and multivariate methods for radiobiological dosimetry. *Radiat Res*. (2010) 173:225-37. doi: 10.1667/RR1836.1

23. Matthews Q, Jirasek A, Lum JJ, Brolo AG. Biochemical signatures of *in vitro* radiation response in human lung, breast and prostate tumour cells observed with Raman spectroscopy. *Phys Med Biol*. (2011) 56:6839–55. doi: 10.1088/0031-9155/56/21/006

24. Matthews Q, Isabelle M, Harder SJ, Smazynski J, Beckham W, Brolo AG, et al. Radiation-induced glycogen accumulation detected by single cell raman spectroscopy is associated with radioresistance that can be reversed by metformin. *PloS One*. (2015) 10. doi: 10.1371/journal.pone.0135356

25. Harder SJ, Isabelle M, Devorkin L, Smazynski J, Beckham W, Brolo AG, et al. Raman spectroscopy identifies radiation response in human non- small cell lung cancer xenografts. *Sci Rep*. (2016) 6:1–10. doi: 10.1038/srep21006

26. Fuentes AM, Narayan A, Milligan K, Lum JJ, Brolo AG, Andrews JL, et al. Raman spectroscopy and convolutional neural networks for monitoring biochemical radiation response in breast tumour xenografts. *Sci Rep*. (2023) 13. doi: 10.1038/s41598-023-28479-2

27. Ali-Adeeb RN, Shreeves P, Deng X, Milligan K, Brolo AG, Lum JJ, et al. Raman microspectroscopy and machine learning for use in identifying radiation-induced lung toxicity. *PloS One*. (2022) 17:e0279739. doi: 10.1371/journal.pone.0279739

28. Cullen D, Bryant J, Maguire A, Medipally D, McClean B, Shields L, et al. Raman spectroscopy of lymphocytes for the identification of prostate cancer patients with late radiation toxicity following radiotherapy. *Transl Biophotonics*. (2020) 2:e201900035. doi: 10.1002/tbio.201900035

29. Nguyen TNQ, Maguire A, Mooney C, Jackson N, Lynam-Lennon N, Weldon V, et al. Prediction of pathological response to neo-adjuvant chemoradiotherapy for oesophageal cancer using vibrational spectroscopy. *Transl Biophotonics*. (2020) 3(1): e202000014. doi: 10.1002/tbio.202000014

30. Esposito R, Portaccio M, Meschini R, Delfino I, Lepore M. Monitoring biochemical changes of neuroblastoma cells in early stages after X-ray exposure by using fourier-transform infrared spectroscopy. *Sensors*. (2024) 24:7459. doi: 10.3390/s24237459

31. Butler HJ, Ashton L, Bird B, Cinque G, Curtis K, Dorney J, et al. Using Raman spectroscopy to characterize biological materials. *Nat Protoc*. (2016) 11:664–87. doi: 10.1038/nprot.2016.036

32. Meade AD, Howe O, Unterreiner V, Sockalingum GD, Byrne HJ, Lyng FM. Vibrational spectroscopy in sensing radiobiological effects: Analyses of targeted and non-targeted effects in human keratinocytes. *Faraday Discuss*. (2016) 187:213-34. doi: 10.1039/C5FD00208G

33. Hill IE, Boyd M, Milligan K, Jenkins CA, Sorensen A, Jirasek A, et al. Understanding radiation response and cell cycle variation in brain tumour cells using Raman spectroscopy. *Analyst*. (2023) 148:2594–608. doi: 10.1039/D3AN00121K

34. Monaghan JF, Byrne HJ, Lyng FM, Meade AD. Radiobiological applications of vibrational spectroscopy: A review of analyses of ionising radiation effects in biology and medicine. *Radiation*. (2024) 4:276–308. doi: 10.3390/radiation4030022

35. Ainsbury EA, Abrantes AM, Baatout S, Baeyens A, Botelho MF, Frey B, et al. Individual Radiation Sensitivity and Biomarkers: Molecular Radiation Biology. In: Baatout S, editor. *Radiobiology Textbook*. Springer International Publishing, Cham (2023). p. 387–424. doi: 10.1007/978-3-031-18810-7_7

36. Angèle S, Romestaing P, Moullan N, Vuillaume M, Chapot B, Friesen M, et al. ATMHaplotypes and cellular response to DNA damage: association with breast cancer risk and clinical radiosensitivity. *Cancer Res*. (2003) 63:8717–25.

37. Bryant J, White L, Coen N, Shields L, McClean B, Meade AD, et al. MicroRNA analysis of ATM-deficient cells indicate PTEN and CCDN1 as potential biomarkers of radiation response. *Radiat Res*. (2020) 193:520–30. doi: 10.1667/RR15462.1

38. Medipally DKR, Nguyen TNQ, Bryant J, Untereiner V, Sockalingum GD, Cullen D, et al. Vibrational spectroscopy of liquid biopsies for prostate cancer diagnosis. *Ther Adv Med Oncol*. (2020) 12: 1–23. doi: 10.1177/1758835920918499

39. White L, Shields L, Howe O, Vega Carrascal I, Maguire A, Meade A, et al. A comparison of radiobiological response in cells exposed to low LET radiation with different beam energies. *Radiat Environ Med*. (2020) 9:1–6. doi: /10.51083/radiatenvironmed.9.1_1

40. Maguire A, Vegacarrascal I, White L, McClean B, Howe O, Lyng FM, et al. Analyses of ionizing radiation effects *in vitro* in peripheral blood lymphocytes with Raman spectroscopy. *Radiat Res*. (2015) 183:407–16. doi: 10.1667/RR13891.1

41. Gautam R, Vanga S, Ariese F, Umapathy S. Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Tech Instrum*. (2015) 2:8. doi: 10.1140/epjti/s40485-015-0018-6

42. Kendall C, Stone N, Shepherd N, Geboes K, Warren B, Bennett R, et al. Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. *J Pathol*. (2003) 200:602–9. doi: 10.1002/path.v200:5

43. Monaghan J F, Cullen D, Wynne C, M. Lyng F, D. Meade A. Effect of pre-analytical variables on Raman and FTIR spectral content of lymphocytes. *Analyst*. (2023) 148:5422–34. doi: 10.1039/D3AN00686G

44. Slattery C, Nguyen K, Shields L, Vega-Carrascal I, Singleton S, Lyng FM, et al. Application of advanced non-linear spectral decomposition and regression methods for spectroscopic analysis of targeted and non-targeted irradiation effects in an in-vitro model. *Int J Mol Sci*. (2022) 23. doi: 10.3390/ijms232112986

45. Rafsanjani MR, Muddiman R, Cullen D, Afzal R, Nally F, McCoy C, et al. Data fusion strategies for multi-modal classification of auto-immune dysregulation with vibrational spectroscopy. *Translational Biophotonics: Diagnostics and Therapeutics III Technical Digest Series* (Optica Publishing Group), paper 126270A (2023). doi: 10.1117/12.2670990

46. McInnes L, Healy J, Saul N, Großberger L. UMAP: uniform manifold approximation and projection. *J Open Source Software*. (2018) 3:861. doi: 10.21105/joss.00861

47. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. (2018) 37:38–44. doi: 10.1038/nbt.4314

48. Sigle M, Rohlfing AK, Kenny M, Scheuermann S, Sun N, Graeßner U, et al. Translating genomic tools to Raman spectroscopy analysis enables high-dimensional tissue characterization on molecular resolution. *Nat Commun*. (2023) 14:1–16. doi: 10.1038/s41467-023-41417-0

49. Liu X, Shi L, Zhao Z, Shu J, Min W. VIBRANT: spectral profiling for single-cell drug responses. *Nat Methods*. (2024) 21:501–11. doi: 10.1038/s41592-024-02185-x

50. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotech*. (2019) 37:38–44. doi: 10.1038/nbt.4314