Dicentric dose estimates for patients undergoing radiotherapy in the RTGene study to assess blood dosimetric models and the new Bayesian method for gradient exposure.

Jayne Moquet[1*], Manuel Higueras[2], Ellen Donovan[3], Sue Boyle[4], Stephen Barnard[1], Clare Bricknell[1], Mingzhu Sun[1], Lone Gothard[4], Grainne O'Brien[1], Lourdes Cruz-Garcia[1], Christophe Badie[1], Elizabeth Ainsbury[1], Navita Somaiah[4].

[1] Public Health England, Centre for Radiation, Chemical and Environmental Hazards (PHE-CRCE), Chilton, Didcot, Oxford OX11 0RQ, UK

[2] Basque Center for Applied Mathematics, Alameda de Mazarredo 14, 48009 Bilbao, Spain

[3] Centre for Vision Speech and Signal Processing, University of Surrey, Guildford, GU2 7XH, UK

[4] Institute of Cancer Research (ICR), 15 Cotswold Road, Sutton, London SM2 5NG, UK

[*] Address for correspondence: Public Health England, Centre for Radiation, Chemical and Environmental Hazards, Chilton, Didcot, Oxfordshire, OX11 0RQ, UK; e-mail: jayne.moquet@phe.gov.uk

ABSTRACT

The RTGene study was focused on the development and validation of new transcriptional biomarkers for prediction of individual radiotherapy (RT) patient responses to ionising radiation (IR). In parallel, for validation purposes, the study has incorporated conventional biomarkers of radiation exposure, including the dicentric assay (DCA). Peripheral blood samples were taken with ethical approval and informed consent from a total of 20 patients undergoing external beam RT for breast, lung, gastrointestinal or genitourinary tumours. For the DCA, two samples were taken from each patient: prior to RT and before the last fraction. Blood samples were set up using standard methods for the DCA. All the baseline samples have dicentric frequencies consistent with the expected background for the normal population. For blood taken before the final fraction, all the samples display distributions of aberrations which are indicative of partial body (PB) exposures. Whole body (WB) and PB cytogenetic doses were calculated with reference to Public Health England (PHE) 250 kVp X-ray calibration curve and then compared to the total dose to blood derived using two newly developed blood dosimetric models. Initial comparisons indicate the relationship between these measures of dose looks very promising, with a correlation of 0.860 ($p=0.001$). A new Bayesian zero-inflated Poisson finite mixture method has been applied to the dicentric data and PB dose estimates show no significant difference ($p<0.001$) with those calculated by the contaminated Poisson technique. The next step will be further development and validation in a larger patient group.

INTRODUCTION

Background to the RTGene Project

Biological markers of radiation exposure play a crucial role in the triage of suspected exposed persons following a radiation accident or incident (IAEA, 2011; Kulka et al., 2017). In recent years the gene expression assay has been shown to be a sensitive marker of radiation exposure, with the potential to be used for truly individualised biological dosimetry (Kabacik et al. 2011a and b; Manning et al. 2013). Classic cytogenetic techniques, and in particular the gold standard dicentric assay, have two main disadvantages in mass casualty scenarios: (1) lack of high-throughput and (2) delays of several days between blood sampling and the availability of results (IAEA, 2011). Using blood samples, gene expression analysis can provide valuable information, as there is a window of time (i.e. 12-48 hours) following radiation exposure where specific radiation-responsive genes have linear dose responses (0-5 Gy) (Manning et al. 2013). New technology for gene expression analysis allows direct counting of nucleic acid molecules (DNA, mRNA, miRNA and lncRNA) without the need for enzymatic reaction or amplification steps hence reducing time for data collection (Kabacik et al. 2015) and has been assessed for radiation biodosimetry applications with promising results (Manning et al. 2011). Linearity of the transcriptional dose-response for specific radiation-responsive genes in *ex vivo* exposed human blood samples has recently been demonstrated for the first time, and inter-individual variability in the response after low doses and high doses exposures has been newly assessed (Kabacik et al. 2011a; Manning et al. 2013). The logical next stage for biological development of the gene expression assay was to further validate these new techniques with human blood samples exposed to radiation *in vivo* (Manning et al. 2017, O'Brien et al. 2018). The RTGene

Project was a feasibility study to develop existing knowledge on coding and non-coding transcriptional responses to IR into a useable radiation specific biomarker of exposure and response using blood samples from RT patients. In parallel, for validation purposes, the study included conventional biomarkers of radiation exposure, i.e. chromosome aberrations using the dicentric assay (DCA) and DNA damage using the gamma-H2AX foci assay.

A range of standard RT schedules were chosen for inclusion in this study to provide a wide range of doses for assessment of the gene expression assay alone and in combination with the DCA, to simulate a wide range of potential exposure scenarios. Conventional cytogenetics was chosen for inclusion in the RTGene project, because the DCA is the most widely used and validated biological dosimetry assay, as well as being a standardised technique (IAEA 2011; ISO 2014). Whole body (WB) and partial body (PB) doses can be assessed based on the observed yield of dicentric chromosome aberrations with reference to an appropriate calibration curve (IAEA, 2011). Not only can cytogenetic dose estimates be used to validate the gene expression assay, but they can be compared to the calculated dose to blood during RT from dosimetric models. In addition, the DCA data can be applied to a more sophisticated Bayesian zero-inflated Poisson finite mixture method to calculate PB dose estimates and then compared to the RT data.

Blood dosimetric models

Radiotherapy treatment planning systems produce detailed maps of the predicted radiation dose to be delivered by the treatment units. The radiation dose is focussed on an area outlined on a computed tomography (CT) image set by a radiation oncologist. This region is

referred to as the target volume and is the site of the primary tumour, tumour bed or region to which the cancer has spread. Radiotherapy is most commonly delivered using photon radiation. Photon radiation is attenuated, but not stopped, by the body hence non-target normal tissues in the path of the beam receive radiation dose. Non-target tissues/organs of particular concern in the vicinity of the target volume are also outlined by the radiation oncologist so that the radiation dose received by them is minimised in the planning process.

Radiotherapy dose information whilst reasonably accurate in the target volume, (IAEA 2004), provides information only on static objects which have been explicitly delineated in the treatment planning system. It can be used to infer dose to the circulating blood but does not give this directly. For this reason, two simple models were set up to test for correlation with the dicentric dose models. The first model used typical values for the circulation time of blood in humans combined with the time taken for radiotherapy to be delivered; the second model used patient specific data of the volume imaged by CT and the mean dose to this volume plus assumptions about the blood volume at the time of irradiation

The Bayesian zero-inflated Poisson finite mixture method to assess partial body exposure

IR produces damage at a cellular level in humans and as mentioned before, the DCA is a well-established cytogenetic biomarker of radiation exposure. To calibrate the effect of IR, dose-response curves are built, by the irradiation of *in vitro* blood samples to different doses, simulating homogeneous whole body exposure. For this kind of exposure, it is typically assumed that the yield of dicentrics per blood cell is a Poisson number whose intensity is a quadratic function of the dose, $\beta_0 + \beta_1 D + \beta_2 D^2$ (for more details, see chapter 8, IAEA 2011).

Gradient exposures are heterogeneous irradiations where different doses occur in the irradiation field within the individual's body. PB irradiations are those where the dose or doses are not absorbed by the whole body of the individual, i.e. there is a fraction of the body which is non-irradiated.

In a scenario of partial body gradient exposure with k irradiated components, given a sample $y = \{y_1, \dots, y_n\}$ of n chromosomal aberration counts within blood cells, the process of the yield of chromosomal aberrations can be represented by a zero-inflated Poisson finite mixture model whose probability mass function has the form

$$p(y_i|\omega, \lambda) = \omega_0 1_{(y_i=0)} + \sum_{j=1}^{k} \omega_j \, p(y_i|\lambda_j),$$

where ω denotes the proportions (with $\sum_{j=0}^{k} \omega_j = 1$), λ is the vector of Poisson intensities, $p(y_i|\lambda_j) = e^{-\lambda} \lambda^{y_i}/y_i!$ is the Poisson probability of observing y_i for expectation $\lambda_j (> 0)$, i is the index of observations, and $1_{(y_i=0)}$ takes the value 1 if $y_i = 0$ and 0 otherwise. Values λ_j and ω_j represent the yield of chromosomal aberrations and the proportion of scored cells at component j respectively. Value ω_0 is the proportion of extra zeroes and represents the proportion of non-irradiated scored cells.

The doses at each component, D_j, are estimated by matching the yield of aberrations to the fitted dose-response curve, $\beta_0 + \beta_1 D_j + \beta_2 D_j^2 = \lambda_j$. To calculate the irradiated body fractions, F_j, is necessary to rescale the proportion of scored cells by adding at each component the proportion of cells which died because of the irradiation, *i.e.*

$$F_0 = \frac{\omega_0}{\omega_0 + \sum_{l=1}^{k} \omega_l e^{D_l/d_0}}, F_j = \frac{\omega_j e^{D_j/d_0}}{\omega_0 + \sum_{l=1}^{k} \omega_l e^{D_l/d_0}},$$

where $d_0$is the 37% cell survival dose, with experimental evidence to be between 2.7 and 3.5 Gy (IAEA 2011). F_0 represents the fraction of the body non-irradiated and F_j represents the fraction of the body irradiated by dose D_j.

This paper looks at the dicentric dose estimates for patients undergoing radiotherapy enrolled in the RTGene study to assess 1) blood dosimetric models and 2) the Bayesian zero-inflated Poisson finite mixture method for estimating partial body exposure. Additional results from gene expression, radiation induced gamma-H2AX foci and translocation analysis, plus a comparison of the cytogenetic dose estimates with the gene expression data will be the subject of separate papers.

MATERIALS AND METHODS

Patient selection and blood sampling

Eligible volunteers who required external beam RT for breast, lung, gastrointestinal or genitourinary tumours were identified in the Outpatient Department at The Royal Marsden NHS Foundation Trust (RM). Patients were included in the study if: a) they were aged 18 years or older; b) had no previous RT; c) not concurrently receiving chemotherapy or not less than four weeks before RT; d) not concurrently receiving hormone therapy or not less than four weeks before RT; e) written informed consent was given, but could be withdrawn at any time. The study was carried out in accordance with the Declaration of Helsinki (1964), the Research Governance Framework 2nd edition (2005) and the Human Tissue Act (2004). The study was approved by the South Central-Hampshire B Research Ethics Committee (16/SC/0307) and registered with ClinicalTrials.gov (NCT02780375). All relevant information was collected from the participants and extracted from their case notes by the research team at RM and only transferred to PHE for the final analysis. No identifiable information was passed between the two institutions and patient confidentiality was maintained at all times. Heparinized venous blood was collected from a total of 20 volunteers, prior to RT and before the last fraction for the DCA. The 20 volunteers were made up of patients undergoing RT treatment for the following tumour types: breast, endometrial, prostate, lung, oesophageal and colon. Coded samples were dispatched by express courier overnight to PHE.

Dicentric assay

On arrival at the laboratory whole blood was mixed with Minimal Essential Medium (MEM) for the DCA (Sigma-Aldrich, Dorset, UK), supplemented with 10% FBS, 1% phytohaemagglutinin, 100 units/ml penicillin plus 100µg/ml streptomycin 2mM L-glutamine (all from Invitrogen, Paisley, UK). In addition, 5-bromo-2-deoxyuridine (Sigma-Aldrich, Dorset, UK) was added to the DCA cultures at a final concentration of 10 µg/ml. All samples were cultured at 37 °C in a 5% CO_2 humidified atmosphere. After 45 hours (h) Colcemid (Sigma-Aldrich, Dorset, UK) was added to each culture to give a final concentration of 0.2 µg/ml. At 50 h metaphases were harvested by a standard hypotonic treatment in 0.075 M potassium chloride for 7 min at 37°C followed by three changes of 3:1 methanol:acetic acid fixative. Fixed cells were dropped onto clean microscope slides, air dried and stained using the fluorescence plus Giemsa technique. The culture, fixation and staining procedures followed the standard protocol recommended by the International Atomic Energy Agency (IAEA, 2011). A maximum of 1000 first division metaphases per donor for the pre-RT sample and 500 cells or 100 dicentrics for the final sample were scored manually for chromosome aberrations. Dose estimates, based on the number of dicentrics per cell were calculated using Dose Estimate_v5.1 (Ainsbury and Lloyd, 2010) and PHEs standard 250 kVp X-ray calibration curve, with following coefficients $C = 0.0005 \pm 0.0005$, $\alpha = 0.046 \pm 0.005$, $\beta = 0.065 \pm 0.003$ (Lloyd et al. 1975). In addition, the standard 'contaminated Poisson' method to calculate the most likely partial body dose, % of lymphocytes exposed and % of the body exposed was applied (IAEA, 2011).

Simple blood dosimetry models

Model 1 (EDD1) uses the high dose (volume within 95% isodose curve on the radiotherapy treatment plan) as a fraction of an assumed 6 litres of blood for a human with a blood flow rate of 6 litres a minute. A RT irradiation time of 1 minute is assumed:

D_B = dose to blood per fraction = D_F x ($V_{95\%}/V_B$) (Gy)

Where, $V_{95\%}$ = high dose volume which is specific to each patient (cc); V_B = total blood volume, assumed to be 6 litres; D_F = radiotherapy dose per fraction which is specific to each patient (Gy). The main limitations of model 1 are a lack of patient specific blood volume, circulation time and no knowledge of volumes of blood in different organs. The uncertainty of $V_{95\%}$ is of the order of 2% – 3%; that of V_B is of the order of 20%; D_F is a set number (the specified dose prescription) and as such does not have an uncertainty; the blood flow rate and RT irradiation times will have variation of at least 20%. It is likely that D_B calculated using this method will have a minimum uncertainty of at least 20%.

Model 2 (EDD2) estimates a whole body mean dose and assumes the blood receives this. The whole body dose is calculated using the mean dose for the volume of the body covered by the CT planning scan. This is scaled assuming the total body volume is 2.5 times this volume:

D_B = dose to blood per fraction = (D_{PB}/D_F) x 2.5 (Gy)

Where, D_{PB} = mean dose (Gy) of body volume covered by CT scan (specific to each patient); D_F = RT dose (Gy) per fraction which is specific to each patient. The main limitations of model 2 are the use of an estimate of 2.5 for the scaling factor from partial to whole body volume and a lack of knowledge of the amount of blood volume in specific organs. The uncertainty of D_{PB} is of the order of 2% – 3%; D_F is a set parameter and as such there is no

uncertainty associated with it; the factor 2.5 is estimated to have an uncertainty of 30% - 40%. It is likely that D_B calculated using this method will have an uncertainty of around 40%. It is non-trivial to determine the volume of blood and blood flow through specific organs relevant to RT and similarly, partial body dose (EDD2) without whole body imaging information. Investigations are underway into more sophisticated estimates using virtual body phantoms.

Bayesian zero-inflated Poisson finite mixture method

The goal is to estimate the absorbed doses and the irradiated fractions for each irradiated component, assuming a total of k.

For estimating both the doses and the fractions a Bayesian model is proposed, which assumes prior distribution densities for the parameters. The technique proposed here consists in two steps, first to infer the yields and the proportions and then to get the estimation of the doses and the fractions by the formulas stated in the section 'The Bayesian zero-inflated Poisson finite mixture method to assess partial body exposure'.

Given sample y, assuming the observations are independent, the likelihood is the product of the probability of the observations $L(\omega, \lambda|y_i) = \prod_{i=1}^{n} p(y_i|\omega, \lambda)$. It is considered ω and all λ_j independents and it is assumed the following prior structure:

$$\omega \sim \text{Dirichlet}(\vec{1}_{k+1}),$$
$$\lambda_j \sim \mathcal{U}(0, M), \lambda_{j+1} > \lambda_j, M = \max(y).$$

The prior for the proportions is a flat Dirichlet distribution of $k + 1$elements. The ordering constraint of the yields prior is to ensure identifiability. By the Bayes' theorem, the joint posterior distribution of $\{\omega, \lambda\}$ is

$$p(\omega,\lambda|y_i) = \frac{L(\omega,\lambda|y_i)\; p(\omega,\lambda)}{\int L(\omega,\lambda|y_i)\; p(\omega,\lambda)\; d\omega\; d\lambda}$$

where $p(\omega,\lambda)$ is the product of the prior densities of all λ_j and ω. The above joint posterior density has a non-tractable form, so acceptance-rejection sampling is used to simulate it. Let $\hat{L}$ the maximum value of $L(\omega,\lambda|y_i)$, the next steps samples the joint posterior distribution:

1. Generate u from $\mathcal{U}(0,M)$.
2. Generate one random variate for each prior, ω^* and λ^*, all them independent of u.
3. Compute $L^* = L(\omega^*,\lambda^*|y_i)$. If $u < \hat{L}/L^*$, then set $\{\omega^*,\lambda^*\}$ to the joint posterior sampling.
4. While the size of the sample is lower than the desired, go to step 1.

To get the joint posterior distribution of the doses and the fractions, a prior is defined for the calibration coefficients, $\{\beta_0,\beta_1,\beta_2\}$, based on the dose-response curve maximum likelihood estimation, and another for the cell survival dose to be uniform between 2.7 and 3.5 Gy, keeping the independency for all priors. This additional prior structure is defined

$$\beta \sim N\left(\hat{\beta},\hat{\Sigma}\right),$$

$$d_0 \sim \mathcal{U}(2.7,3.5),$$

so the following steps are included in the previous algorithm after step 3 if the condition is met:

a) Generate one random variable for the new priors: β^* and d_0^*.

b) Calculate a new sample for the doses by solving $\beta_0 + \beta_1 D_j^* + \beta_2 {D_j^*}^2 = \lambda_j^*$.

c) Calculate the fractions from

$$F_0^* = \frac{\omega_0^*}{\omega_0^* + \sum_{l=1}^{k} \omega_l^* e^{D_l^*/d_0^*}}, F_j^* = \frac{\omega_l^* e^{D_j^*/d_0^*}}{\omega_0^* + \sum_{l=1}^{k} \omega_l^* e^{D_l^*/d_0^*}}.$$

After this process, samples $\{F, D\}$ represent the joint posterior densities and the posterior marginal densities are represented by each F_j or D_j for the joint sample.

This method was applied to the dicentric data to estimate PB doses assuming 2, 3, 4, 5 and 6 irradiated fractions. Due to computational intensity, the number of simulated draws of the joint posterior densities is decreased as the assumption of the number of irradiated components increases. The simulation size for each scenario was as follows: 10000 for 2; 1000 for 3, 4 and 5; 100 for 6 irradiated fractions. A Bayesian Information Criterion (BIC) value was also calculated for the different scenarios.

Other data analysis

In order to investigate whether there was a statistically significant difference in dose response with cancer type, general linear model analysis of variance (GLM ANOVA) was carried out, with post-hoc testing using Tukey's pairwise comparisons within factors, using Minitab® 17. For comparison of the Bayesian and standard PB method, the doses calculated by each technique were normalise and compared using the standard Student's t-test.

RESULTS

The RT schedules and doses are shown in Table 1 for each of the 20 patient included in the study. RT treatment was given to the breast (5 patients); endometrium (4 patients); prostate (3 patients); lung (5 patients); oesophagus (2 patients) and colon (1 patient).

<Table 1>

All baseline samples have dicentric frequencies consistent with the expected background for the normal population (0 – 2 in 1000) and, as expected, there is no indication of departure from the Poisson distribution so there is no indication of recent whole or partial body exposures in these samples. For the samples taken prior to the final fraction, all samples display distributions of aberrations which are indicative of partial body exposures to some degree, as illustrated in Table 2.

<Table 2>

The BIC values for the different exposure scenarios, assuming PB irradiation, were calculated (data not shown). Lower BIC values indicate a better fit. Following this criterion, a PB irradiation with 2 irradiated components was the best fit for the dicentric data for all patients. The results of the cytogenetic dose estimates (standard and Bayesian methods) and total dose to blood calculated from the two models are given in Table 3.

<Table 3>

Figure 1 compares the total doses to blood during RT calculated using ICR/Royal Marsden (ICR/RM) blood dose models 1 (EDD1) and 2 (EDD2) and the dicentric doses to the WB and PB. As illustrated, the relationship between WB dose and EDD2 gives an R^2 correlation of 0.88 and an F-test p-value of 0.001 for the significance of the relationship, the corresponding values for PB dose and EDD2 are 0.72 and 0.001 respectively. For EDD1, there was no significant linear relationship between the model dose and either WB or PB dose, but the R^2 correlations for the plotted relationships were 0.04 and 0.03, respectively. As the models were only initial indications, equal weighting of each point was applied in this case.

<Figure 1>

PB dose estimates calculated by the standard contaminated Poisson method and the new Bayesian technique were compared using the average body doses. These were calculated as the product of the irradiated fraction and dose for the standard method and the sum of the product of the respective doses and fractions for the Bayesian technique. A t-test on these normalised values showed no significant difference ($p < 0.001$) between doses calculated by the Bayesian and the standard method.

Grouping the results in Table 3 by cancer type, in order to investigate whether there is a difference in dose response, calculated using standard methods for the dicentric data, then applying GLM ANOVA for this factor with post hoc testing, showed the type of cancer had a

significant effect on the WB and PB dose ($p < 0.001$). For calculated WB doses, breast cancer treatments showed lower doses than endometrial, lung, oesophageal and colon cancers (p all < 0.001). Prostate cancer patients received lower doses than endometrial, lung, oesophageal (p all < 0.001) and colon cancers ($p = 0.026$). For calculated PB doses, breast cancer treatments resulted in significantly lower doses than treatment for lung cancer ($p = 0.005$) and for cancer of the oesophagus ($p = 0.013$). In addition, treatment for lung cancer resulted in a significantly higher doses than prostate cancer treatment ($p = 0.029$). Cancer of the oesophagus also resulted in a significantly higher dose than treatment for prostate cancer ($p = 0.037$). No other significant differences were observed.

DISCUSSION

Biomarkers of radiation exposure have been used for biological dose estimation for many years; in particular the DCA has been in use since the mid-1960s (IAEA 2011). Biodosimetry methods have the potential to contribute to epidemiological studies of ionising radiation effects (Pernot et al., 2012; Sotnik et al. 2015; Hall et al., 2017). With the aim to improve the application of RT, the significance of predictive and prognostic biomarkers of response to radiation has also been demonstrated (Kerns et al. 2016; Andreassen et al. 2016; Yang et al. 2017). In addition, some studies using cytogenetic biodosimetry assays have shown they may be considered as a predictor of radiosensitivity to identify patients likely to develop acute / chronic adverse effects from RT (Borgmann et al, 2002; Chua et al. 2011; Beaton et al., 2013), although these studies have been small in scale and not prospectively validated. Gene expression analysis has shown possible potential as a marker of radiosensitivity (Badie et al., 2008; Mayer et al., 2011; Finnon et al., 2012) and as a sensitive biological marker for biological dosimetry (Paul and Amundson, 2008; Kabacik et al., 2011a; Badie et al. 2013; Manning et al., 2013). Despite modern techniques the DCA remains the most specific and standardised method for biological dosimetry (Ricoul et al., 2017) and hence it is the assay best suited to validate the gene expression technique for dose estimation.

RTGene was a feasibility study to develop and further validate the gene expression assay for biodosimetry with human blood samples exposed *in vivo* (Manning et al., 2017; O'Brien et al. 2018) and included conventional biomarkers for additional validation. This has allowed dose estimates based on the dicentric assay to be calculated. As Table 1 shows, the RT schedules and doses for the patients are different and the results of the AVOVA analysis

indicate cancer site has a significant effect on the WB and PB dicentric dose estimates. When the cancer sites were compared further significant differences were observed, with treatment for breast and prostate cancer resulting in significantly lower cytogenetic dose estimates than other groups. With breast and prostate RT the high dose volumes are generally smaller than those in other tumour sites resulting in lower WB and PB doses. Breast in particular, if treated with tangential fields only (as is the case in this study), spares the lung and heart, with most of the dose going through less vascular tissue within the breast. Lung and oesophagus RT would invariably result in doses to highly vascular organs such as lung and heart, and this is reflected in the DCA and blood dosimetric models applied here.

Dicentric dose estimates estimated using standard methods (IAEA 2011) have been compared to the calculated dose to blood derived using two newly developed ICR/RM dosimetric models. To the authors knowledge there are currently no recommended published methods to calculate the dose to circulating blood for RT. EDD1 and EDD2 are relatively simple blood dosimetry models, with a number of limitations, such as no knowledge of the blood volume in specific organs. However, as Figure 1 shows, the relationship between the cytogenetic and the model doses is very promising, especially for EDD2. This implies that despite the models crude nature they may be useful and this initial success will allow further development to take place; for example to take account of lymph nodes in the radiation field.

PB dose estimation, termed the contaminated Poisson method, was first proposed for dicentric data in the late 1960s (Dolphin, 1969) and is still one of the standard methods recommended for biological dosimetry (IAEA, 2011). More recently, it has been suggested

that Bayesian statistical analysis may be more suitable for dicentric data (Ainsbury et al, 2014). A new Bayesian zero-inflated Poisson finite mixture method for estimating PB exposure has been developed with test data from simulated PB irradiations, where unirradiated and *ex vivo* irradiated blood samples were mixed in different proportions (Higueras et al., 2016). Cytogenetic data from the RTGene study has allowed the Bayesian zero-inflated Poisson finite mixture method to be used after *in vivo* irradiation and for a comparison of PB dose estimates calculated by this new approach and the standard contaminated Poisson technique. The Bayesian method has shown the distribution of the radiation-induced damage at a cellular level can be expressed in terms of a gradient exposure, but the number of irradiated fractions is lower than the number of RT procedures. In part this difference may be the result of the fractionated nature of the exposure, with a different sub-set of lymphocytes being irradiated during each fraction. However, the good agreement between the Bayesian and standard technique indicate this new method to calculate PB dose has the potential to provide additional information regarding dose estimates and irradiated fraction for biological dosimetry.

In summary, the results from the RTGene study using a conventional biomarker, the DCA, indicate they can be used to validate future gene expression data. Comparisons between the cytogenetic dose estimates and 1) blood dosimetric models and 2) the new Bayesian method for gradient exposure are very encouraging. This will allow further development of the dosimetric models and demonstrates the new Bayesian method can be used for in vivo exposures. A lot more work is needed, but the next step will be further development and validation in a larger patient group. The RTGene partners will also explore the possibility of

combining the cytogenetic, DNA damage and gene expression data to form a multi-assay panel of biomarkers to inform on individual radiation exposure and effects.

ACKNOWLEDGEMENTS

We thank all the patients and staff who participated in the study from the Royal Marsden NHS Foundation Trust, Sutton, in particular Dr Fiona MacDonald (lung), Dr Alison tree (prostate), Dr Susan Lalonrelle (endometrium), Dr Diana Tait and Dr Shree Bhide (gastrointestinal) for recruiting patients into this study. The work was partly supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Chemical & Radiation Threats & Hazards at Newcastle University in partnership with Public Health England (PHE). The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health or PHE. The multi-panel coding and non-coding transcriptional responses as an indicator of individualised responses to radiation effects in radiation therapy patients – RTGene project received a pilot grant from the Opportunity Funds Management Core of the Centers for Medical Countermeasures against Radiation, National Institute of Allergy and Infectious Diseases; grant number U19AI067773, in collaboration with Columbia University. We acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and ICR. The research by MH was supported by the Basque Government through the BERC 360 2014-2017 and the Spanish Ministry of Economy and Competitiveness MINECO and FEDER: BCAM Severo Ochoa excellence accreditation SEV-2013-0323.

TABLES AND FIGURES

Table 1. Table showing the tumour type and the RT schedule and RT prescribed doses for each patient.

RTGene ID	RT treatment to	RT prescribed target dose (Gy)	Number of RT fractions	RT prescribed target dose per fraction (Gy)
RTG002	Breast (right)	40.05	15	2.67
RTG003	Endometrium	45	25	1.80
RTG004	Breast (left)	40.05	15	2.67
RTG005	Breast (left)	40.05	15	2.67
RTG006	Breast (right)	40.05	15	2.67
RTG007	Endometrium	45	25	1.80
RTG008	Prostate	60	20	3.00
RTG009	Lung	55	20	2.75
RTG010	Lung	55	20	2.75
RTG011	Prostate	60	20	3.00
RTG012	Lung	55	20	2.75
RTG013	Lung	55	20	2.75
RTG014	Lung	55	20	2.75
RTG015	Endometrium	45	25	1.80
RTG016	Endometrium	45	25	1.80
RTG017	Prostate	60	20	3.00
RTG018	Oesophagus	36	12	3.00
RTG019	Breast (both)	40.05	15	2.67
RTG020	Oesophagus	20	5	4.00
RTG021	Colon	40	15	2.67

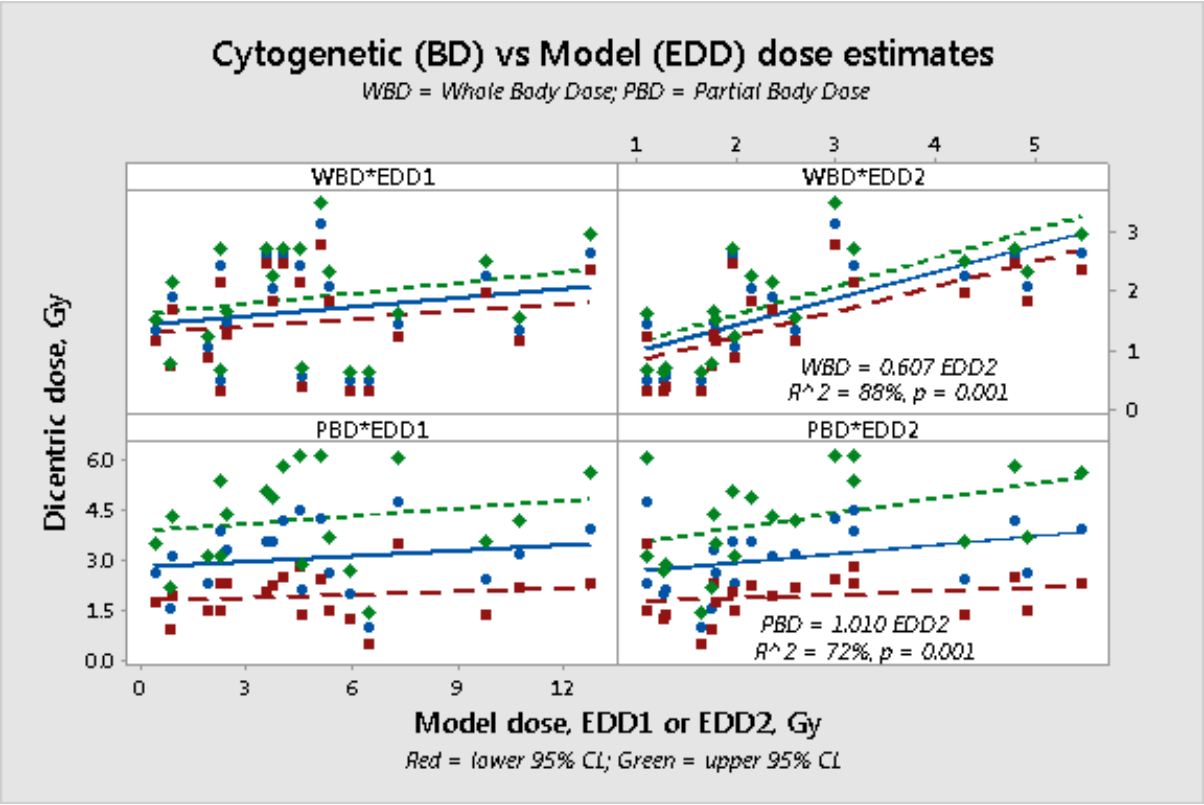
Table 2. Dicentric chromosome aberrations in samples taken prior to the final RT fraction. Cells = Number of peripheral blood lymphocytes scored; Dics = Number of dicentric chromosome aberrations identified; y = yield of dicentrics; var:mean = variance: mean ratio, an indication of departure from Poisson and thus partial body exposure (var:mean for Poisson = 1); SE = standard error of the measurement in the previous column.

RTGene ID	Cells	Dics	Y	SE	var:mean	SE
RTG002	500	20	0.040	0.009	1.360	0.062
RTG003	171	99	0.579	0.058	1.790	0.108
RTG004	500	19	0.038	0.009	1.390	0.062
RTG005	500	23	0.460	0.010	1.300	0.062
RTG006	500	19	0.038	0.009	1.070	0.062
RTG007	233	100	0.429	0.043	1.100	0.092
RTG008	500	36	0.072	0.012	1.150	0.062
RTG009	202	100	0.500	0.050	2.280	0.099
RTG010	488	101	0.207	0.039	1.990	0.064
RTG011	500	60	0.120	0.028	1.350	0.063
RTG012	203	100	0.493	0.091	2.020	0.099
RTG013	309	100	0.323	0.059	1.660	0.080
RTG014	132	103	0.780	0.143	1.830	0.123
RTG015	181	100	0.552	0.102	2.080	0.105
RTG016	264	100	0.379	0.070	1.270	0.087
RTG017	500	90	0.180	0.035	1.820	0.063
RTG018	181	100	0.552	0.102	1.680	0.105
RTG019	500	91	0.182	0.035	1.680	0.063
RTG020	500	99	0.198	0.037	3.010	0.063
RTG021	276	100	0.362	0.050	1.720	0.085

Table 3. Penultimate fraction doses for each participant calculated using blood dosimetric models and the cytogenetic dose estimates calculated by standard and Bayesian methods (95% HPDCI = 95% Highest Posterior Density Credible Interval)

RTGene ID	Blood model doses		Blood cytogenetic dose estimates												
			Whole body		Partial body				Bayesian partial body assuming 2 irradiated fractions (mean values)						
	EDD1 (Gy)	EDD2 (Gy)	Dose (Gy)	SE	Dose (Gy)	SE	% Cells irradiated	% Body irradiated	% Body not irradiated	Dose 1 (Gy)	95% HPDCI	Fraction 1 (%)	Dose 2 (Gy)	95% HPDCI	Fraction 2 (%)
RTG002	2.26	1.08	0.50	0.08	2.34	0.42	8.6	18.3	67.7	1.28	0.00 - 2.77	21.5	2.96	1.45 - 4.89	10.7
RTG003	12.81	5.49	2.65	0.16	4.01	0.84	47.1	79.7	17.9	2.87	0.94 - 4.39	42.9	5.23	3.56 - 7.85	39.1
RTG004	5.94	1.26	0.48	0.08	2.01	0.38	10.7	20.2	58.8	1.04	0.01 - 2.32	32.1	3.27	1.17 - 6.05	9.2
RTG005	4.57	1.28	0.56	0.08	2.14	0.39	11.6	22.5	63.1	1.22	0.001 - 2.57	24.4	2.78	1.14 - 4.62	12.5
RTG006	6.50	1.64	0.48	0.08	0.99	0.24	34.5	43.3	45.9	0.78	0.002 - 1.65	36.7	1.91	0.51 - 4.15	17.5
RTG007	9.81	4.30	2.24	0.14	2.50	0.57	82.5	92.3	8.3	2.14	1.01 - 3.08	49.5	3.24	2.23 - 5.09	42.1
RTG008	0.85	1.75	0.75	0.01	1.60	0.33	29.9	43.6	42.8	1.07	0.03 - 2.02	36.4	2.28	0.97 - 4.25	20.8
RTG009	4.53	3.19	2.43	0.15	4.53	0.86	32.1	71.7	26.8	3.12	0.88 - 4.82	37.1	5.95	4.14 - 8.44	36.1
RTG010	2.45	1.77	1.46	0.09	3.36	0.53	23.4	51.4	41.2	2.35	0.99 - 3.35	45.5	6.48	3.90 - 9.18	13.4
RTG011	1.92	1.98	1.05	0.09	2.34	0.42	25.9	45.4	43.9	1.48	0.18 - 2.72	34.1	3.11	1.75 - 5.11	22.0
RTG012	2.26	3.17	2.42	0.15	3.90	0.78	42.2	75.6	12.3	2.14	1.24 - 3.25	53.7	5.99	4.25 - 7.69	34.1
RTG013	0.91	2.35	1.90	0.12	3.16	0.60	40.7	68.9	17.2	1.89	1.02 - 3.04	57.8	5.05	3.07 - 6.99	25.0
RTG014	5.11	3.00	3.13	0.18	4.33	0.96	55.0	85.9	11.3	2.94	1.30 - 4.59	42.4	5.54	4.05 - 7.49	46.3
RTG015	4.06	4.80	2.58	0.06	4.21	0.85	41.1	76.8	18.3	2.78	1.29 - 4.24	47.0	6.09	4.11 - 8.27	34.8
RTG016	5.33	4.94	2.08	0.13	2.65	0.57	65.6	83.6	12.5	2.06	0.87 - 3.03	53.7	3.80	2.26 - 6.36	33.8
RTG017	0.38	1.78	1.35	0.09	2.68	0.46	30.5	54.2	36.6	2.12	1.38 - 2.82	56.6	8.48	5.16 - 10.85	6.8
RTG018	3.56	1.95	2.58	0.06	3.60	0.78	54.8	82.1	13.0	2.66	1.43 - 3.85	57.4	5.68	3.37 - 8.51	29.6
RTG019	10.77	2.58	1.35	0.09	3.22	0.51	22.1	48.3	45.0	2.02	0.28 - 3.53	30.0	4.03	2.68 - 5.92	25.0
RTG020	7.32	1.08	1.43	0.09	4.82	0.66	11.4	43.5	45.4	1.98	0.51 - 3.59	29.4	6.67	5.31 - 8.26	25.2
RTG021	3.73	2.14	2.03	0.10	3.60	0.67	35.9	68.0	27.4	2.35	0.62 - 3.94	37.0	4.54	3.13 - 6.59	35.7

Figure 1. Penultimate fraction doses to blood during RT calculated using models 1 (EDD1) and 2 (EDD2) and doses to the WB and PB cytogenetic doses, calculated using the standard contaminated Poisson methodology to separate exposed and unexposed fractions in PB exposures.



REFERENCES

Ainsbury, E. and Lloyd, D. Dose estimation software for radiation biodosimetry. Health Phys. 2010. 98(2):290-295.

Ainsbury, E.A., Vinnikov, V.A., Puig, P., Higueras, M., Maznyk, N.A., Lloyd, D.C. And Rothkamm, K. Review of Bayesian statistical analysis methods for cytogenetic radiation biodosimetry, with a practical example. Radiat. Prot. Dosim. 2014. 162(3):185-196. doi:10.1093/rpd/nct301

Andreassen, C.N., Rosenstein, B., Kerns, S., Ostrer, H., De Ruysscher, D., Cesaretti, J., Barnett, G.C., Dunning, A.M., Dorling, L., West, C.M.L., et al. Individual patient data meta-analysis shows a significant association between ATMrs1801516 SNP and toxicity after radiotherapy in 5,456 breast and prostate cancer patients. Radiother. Oncol. 2016. 121(3):431-439.

Badie, C., Dziwura, S., Raffy, C., Tsigani, T., Alsbeih, G., Moody, J., Finnion, P., Levine, E., Scott, D. and Bouffler, S. Aberrant CDKN1A transcriptional response associates with abnormal sensitivity to radiation treatment. Br. J. Cancer. 2008. 98(11):1845-1851. doi: 10.1038/sj.bjc.6604381.

Beaton, L.A., Marro, L., Samiee, S., Malone, S., Grimes, S., Malone, K. and Wilkins, R.C. Investigating chromosome damage using fluorescent *in situ* hybridization to identify

biomarkers of radiosensitivity in prostate cancer patients. Int. J. Radiat. Biol. 2013. 89(12):1087-1093. doi: 10.3109/09553002.2013.825060.

Borgmann, K., Roper, B., El-Awady, R.A., Brackrock, S., Bigalke, M., Dörk, T, at al. Indicators of late normal tissue responses after radiotherapy for head and neck cancer: fibroblasts, lymphocytes, genetics, DNA repair and chromosome aberrations. Radiother. Oncol. 2002. 64:141-152. doi: 10.1016/S0167-8140(02)00167-6.

Chua, M.L., Somaiah, N., A'Hern, R., Davies, S., Gothard, L., Yarnold, J. and Rothkamm, K. Residual DNA and chromosomal damage in ex vivo irradiated blood lymphocytes correlated with normal tissue response to breast radiotherapy. Radiother. Oncol. 2011. 99(3):362-366. doi: 10.1016/j.radonc.2011.05.071

Dolphin, G.W. Biological dosimetry with particular reference to chromosome aberration analysis. A review of methods. Handling of Radiation Accidents. Proc. Int. Symp. Vienna 1969. IAEA, Vienna. 1969.

Finnon, P., Kabacik, S., MacKay, A., Raffy, C., A'Hern, R., Owen, R., Badie, C., Yarnold, J. and Bouffler, S. Correlation of in vitro lymphocyte radiosensitivity and gene expression with late normal tissue reactions following curative radiotherapy for breast cancer. Radiother. Oncol. 2012. 105(3):329-36.

Hall, J., Jeggo, P., West, C., Gomalka, M., Quintens, R., Badie, C., Laurent, O., Aerts, A., Anastasov, A., Azimzadeh, O., Azizova, T., Baatout, S., Baselet, B., Benotmane, M.A.,

Blanchardon, E., Guéguen, Y., Haghdoost, S., Harms-Ringdahl, M., Hess, J., Kreuzer, M., Laurier, D., Macaeva, e., Manning, G., Pernot, E., Ravanat, J-L., Sabatier, L., Tack, K., Tapio, S., Zitzelsberger, H. and Cardis E. Ionizing radiation biomarkers in epidemiological studies – An update. Mut. Res. Reviews. 2017. 771:59-84. doi: 10.1016/j.mrrev.2017.01.001

Higueras, M., Puig, P., Ainsbury, E.A., Vinnikov, V.A., and Rothkamm, K. A new Bayesian model applied to cytogenetic partial body irradiation estimation. Radiat. Prot. Dosim. 2016. 168:330-336. doi:10.1093/rpd/ncv356.

International Atomic Energy Agency (IAEA). Technical Report Series 430: Commissioning and quality assurance of computerized planning systems for radiation treatment of cancer. Vienna: IAEA. 2004.

International Atomic Energy Agency (IAEA). Cytogenetic Dosimetry: Applications in preparedness for and response to radiation emergencies. Vienna: IAEA. 2011.

International Organization for Standardization. ISO 19238: radiation protection – performance criteria for service laboratories performing biological dosimetry by cytogenetics. 2014. Geneva, Suisse:ISO

Kabacik, S., Mackay A., Tamber, N., Manning, G., Finnon, P., Paillier, F., Ashworth, A., Bouffler, S. and Badie, C. Gene expression following ionising radiation: identification of biomarkers for dose estimation and prediction of individual response. Int. J. Radiat. Biol. 2011a. 87(2):115-129. doi:10.3109/09553002.2010.519424

Kabacik, S., Ortega-Molina, A., Efeyan, A., Finnon, P., Bouffler, S., Serrano M. and Badie, C. A minimally invasive assay for individual assessment of the ATM/CHEK2/p53 pathway activity. Cell Cycle. 2011b. 10(7):1152-1161.

Kabacik, S., Manning, G., Raffy, C., Bouffler, S. and Badie, C. Time, dose and ataxia telangiectasia mutated (ATM) status dependency of coding and noncoding RNA expression after ionizing radiation exposure. Radiat. Res. 2015. 183(3):325-337. doi:1667/RR13876.1

Kerns, S.L., Dorling, L., Fachal, L., Bentzen, S., Pharoah, P.D.P., Barnes, D.R., Gómez-Caamaño, A., Carballo, M. Dearnaley, D.P., et al. Meta-analysis of genome wide association studies identifies genetic markers of late toxicity following radiotherapy for prostate cancer. 2016. EBioMedicine. 10:150-163. doi:10.1016/jebiom.2016.07.022

Kulka, U., Abend, M., Ainsbury, E., Badie, C., Barquinero, J.F., Barrios, L., et al., RENEB – running the European Network of biological dosimetry and physical retrospective dosimetry. Int. J. Radiat. Biol. 2017. 93(1):2-14.

Lloyd, D.C., Purrott, R.J., Dolphin, G.W., Bolton, D., Edwards, A.A. and Corp, M.J. The relationship between chromosome aberrations and low LET radiation dose in human lymphocytes. Int. J. Radiat. Biol. 1975. 28:75-90.

Manning, G., Kabacik, S., Finnon, P., Paillier, F., Bouffler, S. and Badie C. Assessing a new gene expression analysis technique for radiation biodosimetry applications. Radiat. Meas. 2011. 46(9):1014-1018.

Manning, G., Kabacik, S., Finnon, P., Bouffler, S. and Badie C. High and low dose responses of transcriptional biomarkers in *ex vivo* X-irradiated human blood. Int. J. Radiat. Biol. 2013. 89(7): p. 512-522.

Manning, G., Tichý, A., Sirák, I. and Badie, C. Radiotherapy-associated long-term modification of expression of the inflammatory biomarker genes ARG1, BCL2L1 and MYC. Front. Immunol. 2017. 8:412. doi: 10.3389/fimmu.2017.00412

Mayer, C., Popanda, O., Greve, B., Fritz, E., Illig, T., Eckardt-Schupp, F., Gomolka, M., Brenner, A. and Schmezer, P. A radiation-induced gene expression signature as a tool to predict acute radiotherapy-induced adverse side effects. Cancer Lett. 2011. 302(1):20-28. doi:10.1016/j.canlet.2010.12.006

O'Brien, G., Cruz-Garcia, L., Majewski, M., Grepl, J., Abend, M., Port., M., Tichý, A., Sirák, I., Malkova, A. et al. FDXR is a biomarker of radiation exposure in vivo. Sci. Rep. 2018. 8:684. doi: 10.1038/s41598-017019043-w

Paul, S. and Amundson, S.A. Development of gene expression signatures for practical radiation biodosimetry. Int. J. Radiat. Oncol. Biol. Phys. 2008. 71(4):1236-1244.

Pernot, E., Hall, J., Baatout, S., Benotmane, M.A., Blanchardon, E., Bouffler, S., Saghire, H., Gomalka, M., Guertler, A., Harms-Ringdahl, M., Jeggo, P., Kreuzer, M., Laurier, D., Lindholm, C., Mkacher, R., Quintens, R., Rothkamm, K., Sabatier, L., Tapio, S., Vathaire, F. and Cardis, E. Ionizing radiation biomarkers for potential use in epidemiological studies. Mut. Res. 2012. 751:258-286.

Ricoul, M., Gnana-Sekaran, T., Piqueret-Stephan, L. and Sabatier, L. Cytogenetics for biological dosimetry. In: Wan T. (eds) Cancer Cytogenetics. Methods in Molecular Biology 2017. 1541:189-208. Humana Press, New York, NY. doi: 10.1007/978-1-4939-6703-2_17

Sotnik, N.V., Azizova, T.V., Darroudi, F., Ainsbury, E.A., Moquet, J.E., Fomina, J. Verification by the FISH translocation assay of historic doses to Mayak workers from external gamma radiation. Radiat. Environ. Biophys. 2015. 54:445-451. doi: 10.1007/s00411-015-0614-5.

Yang, L., Taylor, J., Eustace, A., Irlam J.J., Denley, H., Hoskin, P.J., Alsner, J., Buffa, F.M., Harris, A.L., Choudhury, A. and West, C.M.L. A gene signature for selecting benefit from hypoxia modification of radiotherapy for high-risk bladder cancer patients. Clin. Cancer Res. 2017. 23(16):4761-4768. doi:10.1158/1078-0432.CCR-17-0038