Dose estimation with uncertainty quantification from the γ–H2AX assay

Jochen Einbeck[1], **Rachel Sales**[1], **Manuel Higueras**[2], **Elizabeth Ainsbury**[3], **Stephen Barnard**[3]

[1] University of Durham, Department of Mathematical Sciences
[2] Basque Centre for Applied Mathematics
[3] Public Health England, Centre for Radiation, Chemical and Environmental Hazards

jochen.einbeck@durham.ac.uk

Abstract

Over recent years, the suitability of the gamma-H2AX foci assay as a biomarker for ionizing radiation has been clearly established in principle. However, dose estimation and uncertainty quantification from this assay requires special care due to intra-individual, inter-individual and inter-laboratory variation. This contribution discusses adequate statistical methodology and presents a web applet.

Keywords: Ionizing radiation, retrospective dosimetry, biomarker, γ–H2AX foci, calibration curve

1. Introduction

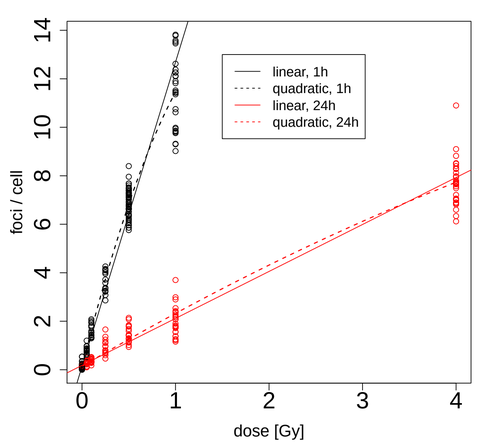
The current "gold–standard" in biodosimetry, based on counts of chromosomal aberrations (dicentrics), is well supported by a solid body of literature including the IAEA manual [1], as well as software (DoseEstimate, CABAS,...). However, the dicentric assay is work– and time–intensive with total global capacity at $<$3000 samples a week [2], which would be insufficient in the case of a large scale accident.

Alternative biomarkers based on proteins, which are cheaper and allow for larger throughput, have recently been developed. Our interest lies in the H2AX–histone, which reacts to radiation–induced DNA double-strand breaks with 'phosphorylation'. The resulting foci can be counted (manually or automated) e.g. by immunofluorescence microscopy. While the suitability of this assay as a radiation biomarker has been documented in the radiological literature in many publications between 2010 and 2015, there have been few systematic efforts at underpinning these developments with statistical methodology which would allow reliable dose estimation. Adaptations of 'dicentric' methodology or software are bound to perform unsatisfactorily since the mechanisms which drive the dose–response relationship are very different for both assays. Specifically, the H2AX foci counts feature much larger (intra-individual, inter-individual, inter-laboratory,...) variation than the dicentric counts.

We introduce methodology which takes these uncertainties into account. Two steps are distinguished in this process: Firstly, the construction of the dose–response curve from in vitro laboratory data, and secondly, the estimation of radiation dose, using the calibration curve, for a new sample of foci counts, from, say, a potentially exposed individual. Due to the count data character of the response, standard least squares regression is not adequate, and instead a quasi-Poisson modelling approach is taken. Dose estimation is carried out through inverse regression, where uncertainties can be decomposed into different sources via the delta method. We present a ready–to–use web applet which implements the developed techniques.

2. Calibration curves

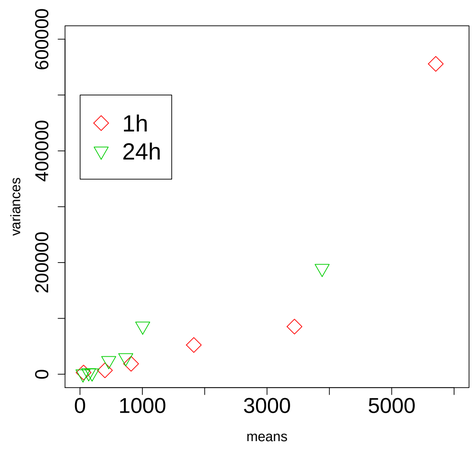
We consider a data set consisting of a total of 339 foci/cell measurements ('yields') following ex–vivo X–radiation of blood taken from 32 individuals (staff volunteers of Public Health England). The design dose points were taken to be 0, 0.05, 0.1, 0.25, 0.5, 1, and 4Gy, and yields (out of $n = 500$ cells) were recorded $1h$ and $24h$ after exposure. The data are displayed below, already including fitted linear and quadratic calibration curves.



It is observed that, for both scenarios, the quadratic curves indicate some small degree of saturation. While this quadratic effect is statistically significant, we found that overall it mainly adds undesired variance to the resulting dose estimates. Hence, we do not consider the quadratic model further, and decide to relate the yields y_{ij} linearly to dose x_i, via

$$\mu_i \equiv E(y_{ij}|x_i) = A + Bx_i. \qquad (1)$$

The key question is which response distribution to use. Since the data come in the form of counts, a Poisson distribution appears adequate. However, plotting means of foci counts per dose point against their variances, we see that the variances are about 60 times higher than the means, violating the equidispersion assumption of the Poisson model.



We proceed, hence, with a quasi-Poisson regression with estimated dispersion $\hat{\phi} \approx 60$. The estimated parameters of this model are identical to the Poisson model, but the parameter uncertainties are adjusted as

$$SE(\hat{A}) = \sqrt{\hat{\phi}}\,\mathrm{SE}_P(\hat{A}) \qquad \mathrm{SE}(\hat{B}) = \sqrt{\hat{\phi}}\,\mathrm{SE}_P(\hat{B}). \qquad (2)$$

where SE_P are the corresponding standard errors from the Poisson model. It is also noted that, via analysis of deviance, we found that inter–individual variation operates on a scale which is not larger than intra–individual variation; hence there is no need to account for these two sources of variation separately, and they can be jointly accounted for by the dispersion parameter.

3. Dose estimation

Assume a blood sample has been taken from a potentially exposed individual, and a number n_* of cells of this sample have been examined for H2AX foci. One will usually have $n_* < n$; for the purpose of triage typically $n_* \approx 50$ [2]. These n_* cells deliver a total foci count Y_* and hence a yield $y_* = Y_*/n_*$. The model above motivates the dose estimator

$$x_* = \frac{y_* - \hat{A}}{\hat{B}}. \qquad (3)$$

The uncertainties can be decomposed via the delta–method. Omitting covariance terms ('MULTIBIODOSE simplification', [3]) and using $\mathrm{SE}(y_*) = \sqrt{\hat{\phi} y_*/n_*}$, one has

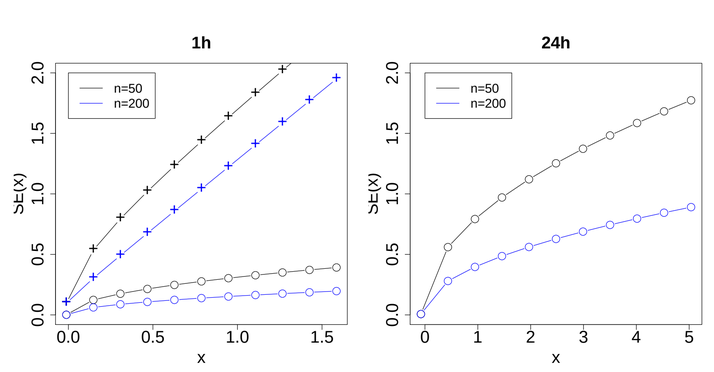
$$\begin{aligned} &SE^2(x_*) \\ &= \left(\frac{\partial x_*}{\partial \hat{A}}\right)^2 SE^2(\hat{A}) + \left(\frac{\partial x_*}{\partial \hat{B}}\right)^2 SE^2(\hat{B}) + \left(\frac{\partial x_*}{\partial y_*}\right)^2 SE^2(y_*) \\ &= \frac{1}{\hat{B}^2} SE^2(\hat{A}) + \frac{(y_* - \hat{A})^2}{\hat{B}^4} SE^2(\hat{B}) + \frac{1}{\hat{B}^2}\frac{\hat{\phi} y_*}{n_*}. \end{aligned}$$

4. Reference samples

While the standard errors derived above account for inter– and intra-individual variation as well as measurement and random error *with mean 0* around the calibration curve, they do not account for systematic deviations (bias) from this curve, for instance by different technology used in laboratories, different scorers, etc. It has therefore been proposed to produce *reference samples* [3], which are irradiated at *known* doses and are used to validate the calibration curve. Typically one will have two such reference samples, a control sample taken at zero dose and a positive sample taken at, say, $1.5Gy$. We omit technicalities of this procedure but just lay out our strategy:

- Check whether the reference samples are within 95% prediction intervals around the estimated calibration curve.
  - If so, the curve is validated.
  - Otherwise, construct a new calibration curve from the reference data (naturally, with larger standard errors).
- Use the chosen calibration curve to estimate the dose according to Section 3.

The figure below gives plots of $\mathrm{SE}(x_*)$ versus dose estimates x_*, for $1h$ (left) or $24h$ (right) after exposure, for $n_* = 50$ (black) or 200 (blue) available cells, respectively. Measurements displayed through circles $\circ$ assume that the given calibration curve is true, while those symbolized through a + symbol assume that it is rejected and hence a reference curve had been used. The yields corresponding to the circles/crosses along the curves are $0, 2, 4, \ldots, 20$ for the 1h data and $0, 1, 2, \ldots, 10$ for the 24h data.

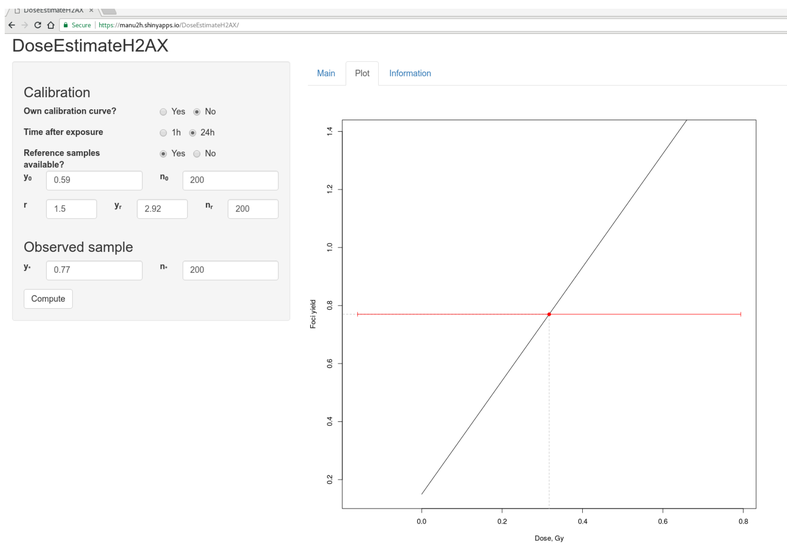


5. Application and web applet

A web applet has been made available at

https://manu2h.shinyapps.io/DoseEstimateH2AX/

which implements this methodology. We produce below an illustration of this applet, using yields obtained at PHE from a different experiment but the same type of radiation.



Several options for the user, depending on the availability of (own) calibration curves, reference samples, standard errors and dispersion estimates are available. Standard errors of dose estimates are adjusted as appropriate. If the required information is not fully available then dose estimates are still produced, but appropriate warning messages are provided.

References

[1] IAEA (2011) *Cytogenetic dosimetry: applications in preparedness for and response to radiation emergencies*. International Atomic Energy Agency.

[2] Rothkamm K, Barnard S, Ainsbury EA, Al–Hafidh J, Barquinero JF, Lindholm C, Moquet J, Perälä M, Roch-Lefèvre S, Scherthan H, Thierens H, Vral A, Vandersickel V (2013). Manual versus automated γ–H2AX foci analysis across five European laboratories: can this assay be used for rapid biodosimetry in a large scale radiation accident? *Mutation Research* **756**, 170–173.

[3] Ainsbury EA, Higueras M, Puig P, Einbeck J, Samaga D, Barquinero JF et al. (2017) Uncertainty of fast biological radiation dose assessment for emergency response scenarios. *International Journal of Radiation Biology* **93**, 127–135.