A new Bayesian model for partial body irradiation dose estimation

M. Higueras[1,2,*], *P. Puig*[2], *E. A. Ainsbury*[1], *K. Rothkamm*[1], *V. V. Vinnikov*[3]

[1]Biological Effects Department, Public Health England; [2]Departament de Matemàtiques, Universitat Autònoma de Barcelona; [3]Grigoriev Institute for Medical Radiology of the National Academy of Medical Science of Ukraine; *Manuel.Higueras-Hernaez@phe.gov.uk

Abstract

A new zero–inflated Poisson model is introduced for cytogenetic estimation of partial body radiation dose and fraction of the body irradiated. Bayes factors are also presented to detect if a sample of chromosomal aberrations in blood cells results from partial or whole body irradiation. One example of simulated cytogenetic exposure is shown to demonstrate the usefulness of this methodology in cytogenetic biological dosimetry.

1. Introduction

The main goal of biological dosimetry is the estimation of the radiation dose received by an exposed individual, in scenarios such as radiation accidents or in radiotherapy settings. Radiation exposure produces breaks in the chromosomal DNA, and the resulting fragments can be repaired in different patterns from their original arrangement. Consequently, the numbers of chromosome aberrations including dicentrics and centric rings increase with the amount of absorbed radiation and are a reliable and very well established biomarker of dose exposure. The estimation of the dose received by an individual requires dose–effect calibration curves, which are produced by exposing peripheral blood lymphocytes to a range of doses, simulating whole body irradiation. The manual of the International Atomic Energy Agency [1] describes the standards for these calibration experiments.

The construction of a calibration curve starts with the irradiation of blood samples from a healthy donor with different doses. Next, the counts of observed chromosomal aberrations are recorded. It is typically assumed that the number of chromosomal aberrations per cell follows a Poisson distribution with a population mean which is a linear–quadratic function of the dose. The set of parameters of this regression model is usually estimated by maximum likelihood, recording the maximum likelihood estimator (MLE) and its variance–covariance matrix. Thus, for an irradiated patient, a blood sample is taken and several tens to approximately one thousand lymphocytes are scored to obtain the counts of chromosomal aberrations. The established approach to infer the absorbed dose and its confidence limits is to use the classical inverse regression method described as a standard procedure in [1].

The mathematical distribution of chromosomal aberration in partial body irradiation (PBI) scenarios is different from Poisson. The non irradiated cells represent an extra amount of zero counts in comparison to the distribution of chromosomal aberrations following whole body irradiation. This proportion of extra zeros can be described by the so–called zero inflated models. The zero counts of aberrations in this zero inflated process result of a mixture of cells with zero

aberrations from the irradiated population and extra zeros which represent the non irradiated cells. Zero–inflated count models provide one method to account for the excess zeros in the data by modelling the data as a mixture of two distributions: a distribution taking a single value at zero and a count distribution, in this case Poisson.

Here, a new cytogenetic method to estimate the absorbed dose and the fraction of the body irradiated in PBI scenarios is presented, based on previous identification of whether the body has been partially or wholly irradiated.

2. Bayesian tools

Assuming the test (patient) data, yields of aberrations, $y = \{y_1, y_2, \ldots, y_n\}$, formed by n count data observations to be $\mathrm{ZIP}(\mu, \omega)$ (μ the population mean and ω the proportion of extra zeros) distributed, the likelihood of this sample remains

$$\mathrm{L}(y|\mu,\omega) \propto \mu^s \sum_{j=1}^{n_0} \binom{n_0}{j} \frac{\omega^j (1-\omega)^{n-j}}{\mathrm{e}^{(n-j)\mu}}, \tag{1}$$

where n_0 and s are respectively the sample frequency of zeros and the sum of the total number of chromosomal aberrations.

The Bayes factor is the main Bayesian model comparison tool. Given a dataset y, the probabilities of two different models, in this case ZIP and Poisson, on y are compared. Following [2] this Bayes factor is defined as

$$BF = \frac{n_0!}{(n+1)!} \sum_{j=0}^{n_0} \frac{(n-j)!}{(n_0-j)!} (1 - j/n)^{-(s+1/2)}. \tag{2}$$

If $2 \log BF > 0, 2, 6, 10$ the evidence in support of the ZIP model is 'weak positive', 'positive', 'strong' and 'very strong' respectively.

Here, the probability of extra zeros ω represents the proportion of non irradiated cells, and D the absorbed dose in the blood sample y. The fraction of the body irradiated, F, which is the proportion of the body which has been exposed to radiation, depending on ω and D. Following [1], ω can be expressed in terms of D and F,

$$F = \frac{(1-\omega)\mathrm{e}^{D/d_0}}{\omega + (1-\omega)\mathrm{e}^{D/d_0}} \Rightarrow \omega = \frac{1-F}{F\mathrm{e}^{-D/d_0} - F + 1},$$

where d_0 is the cell survival dose, with experimental evidence to be between 2.7 and 3.5 Gy [1]. Substituting ω the likelihood in (1) results,

$$L(y|\mu, F, d_0) = \mathrm{L}\left(y \,\middle|\, \mu, \frac{1-F}{F\mathrm{e}^{-D/d_0} - F + 1}\right) \propto \sum_{j=1}^{n_0} \binom{n_0}{j} \frac{F^{n-j}(1-F)^j \mu^s}{(F\mathrm{e}^{-D/d_0} - F + 1)^n \mathrm{e}^{(n-j)\mu}}.$$

Let $\mu = \mathrm{f}(D, \beta)$ be the calibration curve, and $\mathrm{v}(D, \hat{\beta}) = \nabla \cdot \hat{\Sigma}_{\hat{\beta}} \cdot \nabla^T$, where ∇ denotes the gradient of $\mathrm{f}(D, \beta)$ with respect to β and $\hat{\Sigma}_{\hat{\beta}}$ is the variance matrix of β, the *mean prior* is

defined for the purposes of inverse regression as in [3], assuming $\mu|D$ Gamma distributed with mean $\mathrm{f}(D,\beta)$ and variance $\mathrm{v}(D,\hat{\beta})$.

Analogously in [3], because the knowledge of μ implies the knowledge of D, then $L(y|\mu,F,d_0) = L(y|D,\mu,F,d_0)$. Therefore, an application of Bayes' theorem shows the expression of the likelihood of D and F for the given the test data y,

$$\mathcal{L}(y|D,F,d_0) \propto \int_0^{+\infty} L(y|\mu,F,d_0)\mathrm{P}(\mu|D)\mathrm{d}\mu = \sum_{j=1}^{n_0} \binom{n_0}{j} \frac{F^{n-j}(1-F)^j}{(F\mathrm{e}^{-D/d_0}-F+1)^n (n-j)^s} \mathrm{P}(\mathcal{X}_j = s|D), \tag{3}$$

where $\mathcal{X}_j$ is random variable negative binomial distributed with mean $(n-j)\mathrm{f}(x,\hat{\beta})$ and variance $(n-j)\mathrm{f}(D,\hat{\beta}) + (n-j)^2\mathrm{v}(D,\hat{\beta})$.

D, F and d_0 are considered independent random variables, i.e. $\mathrm{P}(D,F,d_0) = \mathrm{P}(D)\mathrm{P}(F)\mathrm{P}(d_0)$. For F a beta prior is applied. For the cell survival dose, d_0, a uniform prior $\mathcal{U}(c,d)$ is applied, where $c \geq 2.7$ and $d \leq 3.5$. These least informative choices are applied in case there is no information about F and d_0, $F \sim \mathrm{Beta}(1,1) \equiv \mathcal{U}(0,1)$ and $d_0 \sim \mathcal{U}(2.7,3.5)$. A prior gamma distribution is applied for the absorbed dose D, as in [3]. This *prior information* can be defined through expert judgement or by an empirical Bayes method. The empirical Bayes method is applied using the MLE of x and its standard error from (3), i.e. a gamma distribution with mean $\hat{D}$ and variance $\hat{\sigma}^2_{\hat{D}}$.

Therefore, the joint posterior density remains,

$$\mathrm{P}(D,F,d_0|y) = \frac{\mathcal{L}(y|D,F,d_0)\mathrm{P}(D,F,d_0)}{\displaystyle\int \mathcal{L}(y|D,F,d_0)\mathrm{P}(D,F,d_0)\mathrm{d}D\mathrm{d}F\mathrm{d}d_0}.$$

To calculate this joint posterior density, with a non tractable form, and its marginal densities, the acceptance–rejection method is used to simulate the posterior distribution.

3. Example

In a recent experiment to simulate PBI, unirradiated and irradiated blood at each dose was mixed, e.g. the exposed blood fraction comprised 10% for 2 Gy for Sample 1 in [4], presenting 1043 blood cells with 0 dicentrics plus centric rings (Dic+CR), 16 with 1, and 3 with 2. This means 1043 cells free of Dic+CR and 22 scored Dic+CR in 1062 blood cells; following notation in Expression 1: $n_0 = 1043$, $s = 22$ and $n = 1062$. The Bayes factor value (Equation 2) for this sample gives 'strong' evidence in support to the ZIP assumption, because $6 < 2\log BF = 9.41 < 10$.

The calibration dose response data for Dic+CR is taken from a recent experiment within a high dose range [5]. In this example, the appropriate dose–response curve is a quadratic function of the absorbed dose, $\mathrm{f}(D,\beta) = \beta_2 D^2 + \beta_1 D + \beta_0$.

To define the prior of D, the empirical Bayes method is applied, $\hat{D} = 1.64$ and $\hat{\sigma}^2_{\hat{D}} = 0.51$, then the prior of the absorbed dose is $D \sim \mathrm{Gamma}(5.27, 3.22)$. The less informative priors are taken for $F \sim \mathcal{U}(0,1)$ and $d_0 \sim \mathcal{U}(2.7,3.5)$.

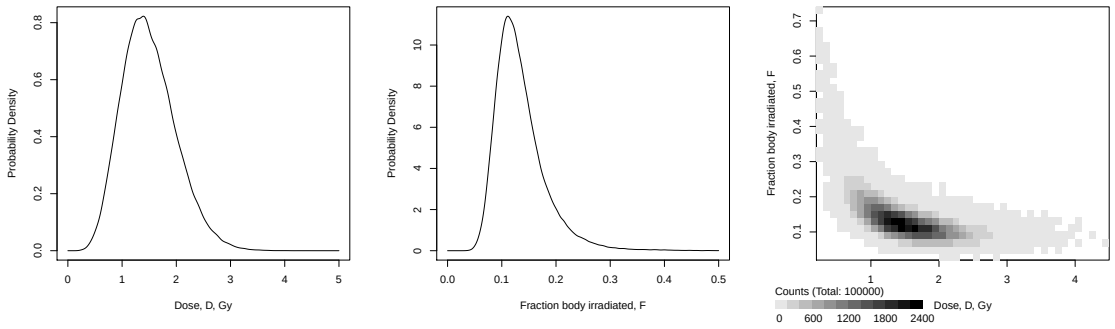

Figure 1: Marginal posterior dose (left) and FBI (center) densities, and histogram of the joint posterior density of (D, F) (right).

Figure 1 show the plots of the posterior densities. The total number of simulations is 100000. These marginal posterior densities return 1.39 Gy as modal dose, 1.50 Gy as expected dose, 0.50 Gy as standard deviation (SD) dose, and (0.66, 2.58) Gy as 95% credible interval (CI) dose; 0.11 as modal F, 0.13 as expected F, 0.05 as SD F, and (0.07, 0.25) as 95% CI F.

In contrast to the classical estimation method the cell survival dose is considered a random variable, while in the classical method is a point value, fixed before the estimation. Future work includes to study the application of informative priors for d_0, and estimation in gradient exposure scenarios.

4. Bibliography

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

[2] Bayarri, M. J., Berger, J. O., Datta and G. S. Objective Bayes testing of Poisson versus inflated Poisson models. *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh* 105–121 (2008), Institute of Mathematical Statistics, see http://projecteuclid.org/euclid.imsc/ 1209398464.

[3] Higueras, M., Puig, P., Ainsbury, E. A. and Rothkamm, K. A new inverse regression model applied to radiation biodosimetry. *Proc. R. Soc. A*, DOI: 10.1098/rspa.2014.0588 (2015).

[4] Higueras, M., Puig, P., Ainsbury, E. A., Rothkamm, K and Vinnikov, V. A. A New Bayesian Model Applied to Cytogenetic Partial Body Irradiation Estimation. *Radiat. Prot. Dosim.*, submitted.

[5] Vinnikov, V. A. and Maznyk, N. A. Cytogenetic dose-response *in vitro* for biological dosimetry after exposure to high doses of gamma-rays. *Radiat. Prot. Dosim.* 154(2), 186–197 (2013).