

radir package: an R implementation for cytogenetic biodosimetry dose estimation

David Morina^{1,2}, Manuel Higuera^{3,4}, Pedro Puig⁴,
Elizabeth A Ainsbury³ and Kai Rothkamm^{3,5}

¹ Centre for Research in Environmental Epidemiology (CREAL), Universitat Pompeu Fabra (UPF), CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona 08003, Spain

² Unitat de Fonaments de l'Anàlisi Econòmica, Departament d'Economia i Història Econòmica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain

³ Centre for Radiation, Chemical and Environmental Hazards, Public Health England, Chilton, Oxfordshire OX11 0RQ, UK

⁴ Departament de Matemàtiques, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain

⁵ University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

E-mail: dmorina@creal.cat

Received 6 February 2015, revised 11 May 2015

Accepted for publication 16 June 2015

Published 10 July 2015



CrossMark


Abstract

The Bayesian framework has been shown to be very useful in cytogenetic dose estimation. This approach allows description of the probability of an event in terms of previous knowledge, e.g. its expectation and/or its uncertainty.

A new R package entitled *radir* (radiation inverse regression) has been implemented with the aim of reproducing a recent Bayesian-type dose estimation methodology. *radir* adopts the method of dose estimation under the Poisson assumption of the responses (the chromosomal aberrations counts) for the required dose-response curve (typically linear or quadratic).

The individual commands are described in detail and relevant examples of the use of the methods and the corresponding *radir* software tools are given. The suitability of this methodology is highlighted and its application encouraged by providing a user-friendly command-type software interface within the R statistical software (version 3.1.1 or higher), which includes a complete manual.

Keywords: cytogenetic dosimetry, Bayesian methods, inverse regression, R software, *radir* package, radiation, dose estimation

 Online supplementary data available from stacks.iop.org/JRP/35/030557/mmedia

(Some figures may appear in colour only in the online journal)

1. Introduction

The classical methods for dose estimation in radiation cytogenetics are well established and described in detail in the manual of the International Atomic Energy Agency (IAEA) [1]. First, calibration data (generally yields of chromosome aberrations in blood lymphocytes) are collected and fitted to a linear or quadratic model, the coefficients of which are then used to calculate doses. The Poisson model is used to describe the uncertainty on the yield of aberrations, and this is combined with uncertainty on the fitted calibration coefficient(s) using standard methodology in order to give the total uncertainty associated with the estimated dose.

In the classical or ‘frequentist’ framework the coefficients of the calibrated dose-response curve are considered ‘fixed’, thus providing an estimate of radiation dose and associated confidence limits using standard likelihood methods. Therefore, assignment of a probability to an event is based solely on the observed frequency of occurrence of the event.

Alternatively, the Bayesian approach considers the parameters, for instance the dose-response curve coefficients in this case, random variables for which previous information could exist. This information could come from previous analysis in the field literature or even from the experts’ opinion/knowledge. The newly collected data is combined with this *prior* information to produce a *posterior* model.

The Bayesian inference uses distributions for all the parameters, leading to an important advantage: the uncertainty of the system is an intrinsic part of the analysis. A review of these methods can be found in [2]. Bayesian methods in cytogenetic biodosimetry give the estimation of the absorbed dose by an individual in a form of random variable distribution, called the calibrative dose density [3].

Some specific software has been developed to fit dose-response curves and to estimate the dose absorbed by an individual; e.g. CABAS [4], DoseEstimate [5] and BioDoser [6]. There is also one recent program, CytoBayesJ [7], which provides some basic software tools for Bayesian analysis of cytogenetic radiation dosimetry data.

In this paper we present a new R statistical software package which implements the Poisson models developed in [8]. These models involve Bayesian-type inverse regression. They use dose-effect calibration curves estimated by the frequentist approach.

This methodology collects the prior information of the yield of chromosomes per cell (the prior population mean) from the dose-effect calibration curve in an univariate parameter by means of the delta method. This prior distribution is assumed normal (under some constraints) or gamma distributed and the calibrative dose density results in terms of the probability functions of the Hermite or negative binomial distributions.

2. The `radir` R software package

The software introduced in this work has been written in the R programming language [9], which is becoming more popular in the cytogenetic biodosimetry context in recent years because of its availability; it can be freely downloaded from <http://cran.r-project.org/> and used on the most common operating systems. In fact, an R script for fitting dose-response curves written by Braselmann was included in [1]. In addition, the `hermite` R package version 1.0.1 [10] has been utilized for the management of the Hermite distribution. The general workflow of `radir` package is summarized in section 2.2.

A video tutorial has been prepared with the aim of helping `radir` users in the installation and general usage of the R statistical software and, in particular, the `radir` package. It can be found in the next open access link, http://polimedia.uab.cat/#v_592.

2.1. Features of *radir* package

In version 1.0, already available from http://cran.r-project.org/src/contrib/radir_1.0.tar.gz, the following tools are included:

- Calculation of the calibrative dose density for a given
 - * expression of the dose-response curve;
 - * hyperparameters set;
 - * estimate of the parameter set;
 - * variance–covariance matrix of the estimation;
 - * total number of cells examined;
 - * number of chromosomal aberrations;
 - * prior distribution of the chromosomal aberration mean: normal or gamma;
 - * prior distribution of the absorbed dose: uniform or gamma;
 - * parameters of the distribution of the dose prior.
- Summary statistics of the calibrative dose density: best estimate, expected value, standard deviation and the 95% highest posterior density (HPD) interval, defined as the shortest range that contains the 95% (or the required percentage) region of the probability density.
- Calculation of the HPD interval for a given credible region.
- Calculation of the probability between two given doses.
- Plots of the
 - * calibrative dose density;
 - * HPD interval for a given credible region;
 - * probability between two given doses;
 - * cumulative dose distribution function.

2.2. *radir* package workflow

The calibrative density is computed explicitly for the Poisson model in [8], which is the most common situation and is also the case covered by the *radir* package. The software takes as inputs laboratory information such as the dose-response curve, the maximum likelihood estimates of its parameters, the variance–covariance matrix and the mean prior distribution, which can be normal or gamma, together with patient information such as the number of cells examined, chromosomal aberrations counts, and the prior dose distribution, which can be uniform or gamma. From this input data, the calibration dose density is calculated and summary statistics, probability between two given doses, high posterior density intervals and several plots can be obtained. This workflow is summarized in figure 1.

2.3. Calibrative dose density calculation

The calculation of the calibrative density is based on the fact that the likelihood function of the data coming from the patient is proportional to a Poisson probability function evaluated at the total number of chromosomal aberrations. Then, the calibrative density remains proportional to the prior dose density multiplied by a probability function of a mixed-Poisson distribution evaluated at the total number of chromosomal aberrations. The nature of the mixed-Poisson distribution depends on the mean prior distribution that acts as the ‘mixing’ distribution. When the mean prior is gamma (the default case for the *radir* package) it leads to a negative binomial distribution, and for the normal density option it leads to the Hermite distribution. The details of all the process and methodology are thoroughly described in [8].

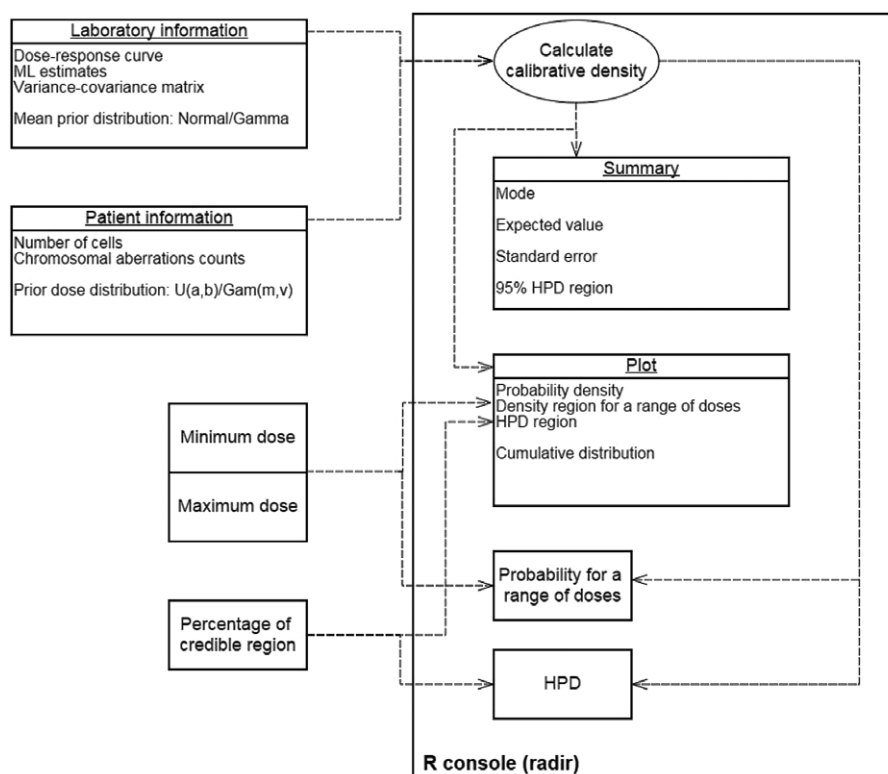


Figure 1. radir package workflow.

The first step is to compute the calibrative dose density, by means of the function `dose.distr()`.

A call to this function might be

```
dose.distr(f, pars, beta, cov, cells, dics, m.prior, d.prior,
  prior.param, stdf)
```

The description of these arguments can be summarized as follows:

- `f`: Dose-response function, as an expression. Must be differentiable in the domain of parameters.
- `pars`: string vector containing the parameters in `f`.
- `beta`: Estimates of the parameters.
- `cov`: Variance-covariance matrix of estimates `beta`.
- `cells`: Patient information: number of cells examined.
- `dics`: Patient information: observed number of aberrations.
- `m.prior`: String containing the prior distribution of the mean. In the current version of the package, it can be `gamma` (the default value) or `normal`.
- `d.prior`: String containing the prior distribution of the dose. In the current version of the package, it can be `gamma` or `uniform` (the default value).
- `prior.param`: Vector of length 2 containing the parameters of the distribution of the dose prior. Its default value is the non-informative prior. If `d.prior` is a gamma distribution, the mean and standard deviation should be given, otherwise the function will return an error.

- `stdf`: Approximated standard deviation factor. This input is useful to control the ends of the calibrative density; i.e. in case the tails of the calibrative dose density are very long this value could be reduced, or *vice versa*. Its default value is 6.

The gamma and normal distributions are the alternatives for the mean prior distribution. In principle, the normal distribution would be the most natural choice because, according to maximum likelihood theory, the mean is asymptotically normal distributed with expectation and variance depending on the dose and the dose-response function. The calibrative density is then proportional to the prior dose density multiplied by a probability function of a Hermite distribution, that is, a Poisson-mixed normal distribution. However, to mix a Poisson with a normal distribution only makes sense when the population mean of the normal distribution is greater than its variance. For this reason, when using the normal mean prior option, the range of doses of the calibrative density could be truncated to the right, and potentially not being enough sensible for the considered sample. Conversely, using the gamma mean prior distribution (the default), mixing a Poisson with a gamma distribution (negative binomial distribution) does not create such problems and the range of doses is not truncated. Moreover, a gamma distribution with a large shape parameter is a good approximation to the normal distribution. Therefore, it is recommended in general to use the gamma mean prior option.

The gamma and uniform distributions are the alternatives for dose prior distributions. The gamma distribution has been used for instance in [11], and is parametrized here in terms of its mean and standard deviation. The uniform distribution is parametrized by its minimum and maximum, and to the knowledge of the authors it has not been previously used for cytogenetic dosimetry, even though is a sensible choice in general dosimetry. The dose prior choice depends on expert opinion and/or the information collected from the irradiation event. If there is no prior information, the most appropriate option is to choose an improper uniform prior defined between zero and infinity. This non-informative option is the default. If limited knowledge about the dose is available, for instance its maximum range, then we could use as a prior a proper uniform distribution defined over zero and the maximum range. More information about the dose (for instance mean and standard deviation), as in the example 3.1, can lead to the use of a gamma prior.

The function output collects the sequence of doses and their respective probability density.

2.4. Statistics summary, credible region, and probability between doses

A summary containing the most relevant information about the estimated doses can be obtained via `summary()`.

This function, when applied to the output of `dose.distr`, gives the most interesting statistics in this context including mode, expected value, standard deviation and 95% HPD credibility interval.

The HPD credible interval for an object of class `dose.radir` can be obtained numerically by means of the function `ci.dose.radir`, with these parameters:

- `object`: an object of class `dose.radir` containing the estimated doses.
- `cr`: credible region size. Its default value is 0.95.

The probability between two doses can be obtained numerically by means of the function `pr.dose.radir`, with parameters

- `object`: an object of class `dose.radir` containing the estimated doses.
- `lod`: lower dose value. Its default value is 0.
- `upd`: upper dose value. Its default value is the maximum dose in `object`.

2.5. Plots

Graphics can be obtained in the standard way by means of `Rplot()` or `lines()` functions.

The `plot` function can also be used to present credible intervals through the argument `ci=TRUE`. The desired credible region size can be fixed using the argument `cr`, which is 0.95 by default. The colour of the shaded credible region is grey by default, but it may be changed by using the argument `col.ci`. For instance, to see the credible region shaded in red, the user should write `col.ci="red"`.

The probability between two doses can be graphically represented by means of the argument `prob=c(d1, d2)`, where `d1` and `d2` are respectively the lower and upper doses considered. The colour of the shaded region is grey by default, but it may be changed by using the argument `col.pr` in the same way as for the parameter `col.ci`.

The distribution function can be plotted as well, using the argument `distr=TRUE` in the `plot` function.

3. Examples

Several examples of use of the `radir` package (some of them introduced in [8]) are described in detail in this section.

3.1. Cobalt-60 gamma-ray irradiation

In [8] the authors consider an example from an *in vitro* cobalt-60 gamma-ray exposure. From the calibration data (table 1 in [8]) the 1.5 Gy row is removed to be inferred later. The model consists of a linear-quadratic dose response without a constant term curve, $\beta_2 x^2 + \beta_1 x$, where x represents the absorbed dose, assuming that the counts of the chromosomal aberrations are Poisson distributed. The specific data used in this example (and in the others) were obtained from real experiments, and are reasonable for a cobalt-60 gamma-ray exposure.

Maximum likelihood estimation provides

$$\hat{\beta}_1 = 3.126 \cdot 10^{-3}, \hat{\beta}_2 = 2.537 \cdot 10^{-2}, \hat{\Sigma}_{\hat{\beta}} = \begin{pmatrix} 7.205 & -3.438 \\ -3.438 & 2.718 \end{pmatrix} \cdot 10^{-6}.$$

The 1.5 Gy sample consisted of 102 observed dicentric in a total of 1811 blood cells. Therefore, the calibrative dose density for this observed data (assuming an improper uniform dose prior) can be calculated with the `radir` package by means of:

```
library(radir)
f <- expression(b1*x+b2*x^2)
pars <- c("b1", "b2")
beta <- c(3.126e-3, 2.537e-2)
cov <- matrix(c(7.205e-06, -3.438e-06, -3.438e-06, 2.718e-06),
             nrow=2)
ex1.a <- dose.distr(f, pars, beta, cov, cells=1811, dics=102,
                  m.prior="normal")
```

The default situation in `dose.distr()` is a gamma mean prior, as

```
ex1.b <- dose.distr(f, pars, beta, cov, cells=1811, dics=102)
```

In [8] the authors consider that, assuming that the real dose is unknown, a reasonable prior dose distribution is gamma with mean 1.75 and standard error 0.375. This can be implemented in the `radir` package by means of

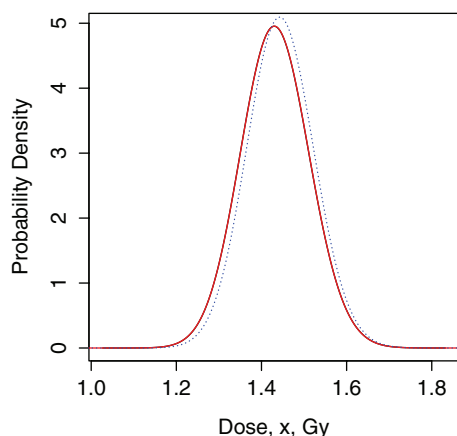


Figure 2. Calibrative densities of the 1.5 Gy test data for a normal mean prior and a $\mathcal{U}(0, \infty)$ dose prior (black, `ex1.a`), for gamma mean priors and a $\mathcal{U}(0, \infty)$ dose prior (red, `ex1.b`) and a gamma dose prior (blue dotted line, `ex1.c`). Note that the black and red lines are indistinguishable (adapted with permission from [8] under CC-BY licence 2015).

```
ex1.c <- dose.distr(f, pars, beta, cov, cells=1811, dics=102,
  d.prior="gamma", prior.param=c(1.75,0.375))
```

Figure 2 shows the plot of the three densities of the estimated dose for the test data. It has been obtained using the `plot()` function on the outputs of function `dose.distr(ex1.a, ex1.b and ex1.c)`, generated from the following code:

```
plot(ex1.a)
lines(ex1.b, col="red")
lines(ex1.c, col="blue", lty=3)
```

It should be observed that these results incorporate the real dose (1.5 Gy) and show the equivalence of both mean priors. Note that the gamma mean prior is moderately more conservative. A summary table of the statistics of the three calibrative densities calculated in this example can be obtained via the `summary()` function. For instance, for the first case:

```
summary(ex1.a)
Mode
-----
1.43
Expected value
-----
1.432
Standard Dev.
-----
0.081
95% CI
-----
(1.275; 1.591)
```

A figure showing the density and the 90% HPD interval (figure 3) can be obtained by means of `plot(ex1.a, ci=T, cr=0.90)`

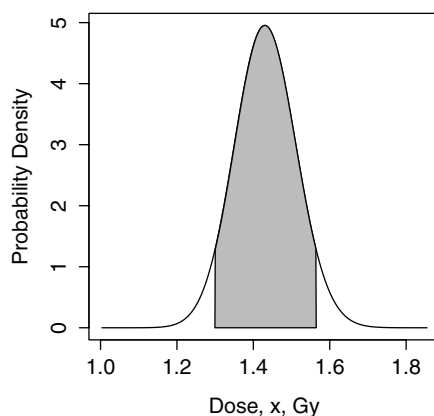


Figure 3. 90% HPD interval of the calibrative density of the 1.5 Gy test data for a normal mean prior and a $\mathcal{U}(0, \infty)$ dose prior.

3.2. Analysis of doses in thyroid cancer patients

Serna *et al* [12] studied chromosomal damage in lymphocytes of thyroid cancer patients after radioiodine treatment. The authors carried out the micronucleus assay in binucleated cells of blood samples from 25 patients 3 d after iodine-131 (3.7 GBq) exposure. The *in vitro* dose-response curve was fitted by a linear-quadratic model, $f(x, \beta) = G\beta_2x^2 + \beta_1x + \beta_0$ according to Poisson’s law, and the estimate of β_0 was not taken into account, because the authors in [12] argued that the intercept could change for each patient. The constant G is the Lea-Catcheside generalized dose-protraction factor, which modifies the quadratic term according to the temporal pattern of exposure; for the *in vitro* assay, $G = 1$. The authors calculated the following parameter estimates ($\hat{\beta}_i \pm SE(\hat{\beta}_i)$),

$$\hat{\beta}_1 = (13.6 \pm 5.5) \cdot 10^{-3}, \hat{\beta}_2 = (3.7 \pm 1.6) \cdot 10^{-2}, \rho = -0.89,$$

where ρ is the correlation coefficient. Taking into account the characteristics of the iodine-131 treatment, the authors in [12] found the factor G to be close to 0.1. Then β_0 , the background of each patient, was estimated by counting the micronuclei of the patient from a blood sample taken before the treatment, information provided in [12]. This leads to the fitted regression model $f(x, \hat{\beta}) = G\hat{\beta}_2x^2 + \hat{\beta}_1x + \hat{\beta}_0$ with a covariance matrix that incorporates the variance of $\hat{\beta}_0$ without correlation with $\hat{\beta}_1$ and $\hat{\beta}_2$. For instance, Patient 1 presented 487 normal cells and 13 cells with just one micronucleus each for a total of 500 cells scored. Before the treatment 5 micronuclei were found in 500 blood cells, thus $\hat{\beta}_0 = (10 \pm 4.472) \cdot 10^{-3}$.

A gamma mean prior is preferred instead of a normal in this case, because the range of doses supported by the normal mean prior is very small, due to mathematical constraints. In [8] the authors show that in that case, the predictive posterior distribution represents the probability of a negative binomial random variable taking a value of 13 counts, with mean $4.810 \cdot 10^{-3}x^2 + 0.177x + 0.130$ and variance $4.326 \cdot 10^{-6}x^4 - 2.647 \cdot 10^{-4}x^3 + 1.008 \cdot 10^{-2}x^2 + 0.177x + 0.133$, for Patient 1. It is possible to define all needed input values for the `radirr` package to analyse the patient data via


```
f <- expression(b0+b1*x+0.1*b2*x^2)
pars <- c("b0", "b1", "b2")
beta <- c(0.01, .0136, .0037)
cov <- matrix(c(1.98e-05, 0, 0, 0, .3121*10^(-4), -.0798*10^(-4), 0,
  -.0798*10^(-4), .0256*10^(-4)), nrow=3)
```

Three calibrative densities have been calculated applying two different proper uniform prior dose distributions, both using information given in [12]. An administered radioiodine activity that produces a blood dose less than 2 Gy is considered safe in the context of medical uses of radiation [13], so one could take a uniform dose prior distribution from 0 to 2 (ex2.u1). On the other hand, the calibration curve was calculated up to a dose of 4.5 Gy, so another uniform dose prior distribution could be from 0 to 4.5 (ex2.u2). An improper uniform prior dose distribution from 0 to ∞ is also applied (ex2.u3). This can be done with the `radir` package by means of

```
ex2.u1 <- dose.distr(f, pars, beta, cov, cells=500, dics=13,
  prior.param=c(0, 2))
ex2.u2 <- dose.distr(f, pars, beta, cov, cells=500, dics=13,
  prior.param=c(0, 4.5))
ex2.u3 <- dose.distr(f, pars, beta, cov, cells=500, dics=13)
```

Table 1 shows the summary results for the 25 patients described in [12]; these results were obtained using a loop that runs the function `dose.distr` for each patient, taking the pre- and the post-radiotherapy information for each of the two uniform prior dose distributions indicated ($U(0,2)$ and an improper uniform distribution). The code that produces these results is provided as supplementary material (stacks.iop.org/JRP/35/030557/mmedia).

3.3. New model for low and high doses

In [14] the authors present a new model for biological dosimetry under a weighted Poisson assumption, where the mean of the underlying Poisson is a Gompertz function of the dose, and the underdispersion level is a linear function of the dose. This leads to a model where the mean of dicentric is

$$f(x, \beta) = \beta_0 e^{-\beta_1 e^{-\beta_2 x}} \left(1 + \frac{\beta_3 x (2\beta_0 e^{-\beta_1 e^{-\beta_2 x}} + 1)}{1 + \beta_3 x (\beta_0^2 (e^{-\beta_1 e^{-\beta_2 x}})^2 + \beta_0 e^{-\beta_1 e^{-\beta_2 x}})} \right). \quad (1)$$

This model is especially useful for high dose exposures.

The *in vitro* irradiation experiment was performed using 10 different doses, from 0 to 25 Gy, and numbers of dicentric in blood lymphocytes were then counted. The models in `radir` are only for the Poisson assumption, so a Poisson model is defined with the dose-response curve defined by expression (1). The maximum likelihood estimation is

$$\hat{\beta}_0 = 8.676, \hat{\beta}_1 = 7.262, \hat{\beta}_2 = 0.230, \hat{\beta}_3 = 2.388,$$

$$\hat{\Sigma}_{\hat{\beta}} = \begin{pmatrix} 0.056 & 0.006 & -0.001 & 0.019 \\ 0.006 & 0.089 & 0.001 & 0.305 \\ -0.001 & 0.001 & 0.000 & 0.003 \\ 0.019 & 0.305 & 0.003 & 1.146 \end{pmatrix}.$$

Table 1. Statistics summary of the calibrative densities of the 25 patients in [12] for $\mathcal{U}(0, 2)$ and $\mathcal{U}(0, \infty)$ dose priors.

Prior dose Patient	$\mathcal{U}(0, 2)$				$\mathcal{U}(0, \infty)$		
	Expected	SD	95% HPD	Mode	Expected	SD	95% HPD
1	1.141	0.481	(0.319, 2.000)	1.140	1.593	0.919	(0.027, 3.362)
2	1.155	0.471	(0.355, 2.000)	1.141	1.588	0.894	(0.079, 3.365)
3	0.759	0.477	(0.000, 1.679)	0.475	0.867	0.637	(0.000, 2.121)
4	1.237	0.444	(0.470, 2.000)	1.279	1.739	0.908	(0.231, 3.599)
5	1.099	0.470	(0.326, 2.000)	1.007	1.434	0.828	(0.057, 3.068)
6	0.847	0.487	(0.000, 1.748)	0.605	1.001	0.693	(0.000, 2.358)
7	1.172	0.459	(0.398, 2.000)	1.143	1.587	0.870	(0.144, 3.343)
8	1.011	0.488	(0.206, 1.977)	0.871	1.290	0.806	(0.000, 2.841)
9	0.842	0.478	(0.000, 1.735)	0.602	0.982	0.671	(0.000, 2.288)
10	1.116	0.457	(0.374, 2.000)	1.008	1.431	0.797	(0.133, 3.041)
11	0.791	0.499	(0.000, 1.729)	0.482	0.949	0.715	(0.000, 2.356)
12	0.452	0.389	(0.000, 1.266)	0.000	0.471	0.430	(0.000, 1.344)
13	0.851	0.495	(0.000, 1.760)	0.608	1.025	0.722	(0.000, 2.433)
14	1.020	0.479	(0.236, 1.976)	0.871	1.280	0.777	(0.000, 2.773)
15	0.580	0.435	(0.000, 1.470)	0.222	0.624	0.515	(0.000, 1.655)
16	0.542	0.437	(0.000, 1.450)	0.000	0.589	0.524	(0.000, 1.653)
17	0.746	0.467	(0.000, 1.655)	0.471	0.840	0.610	(0.000, 2.038)
18	0.607	0.449	(0.000, 1.521)	0.225	0.663	0.544	(0.000, 1.757)
19	0.940	0.470	(0.138, 1.880)	0.734	1.117	0.700	(0.000, 2.465)
20	0.771	0.486	(0.000, 1.699)	0.478	0.899	0.673	(0.000, 2.215)
21	0.771	0.486	(0.000, 1.699)	0.478	0.895	0.663	(0.000, 2.202)
22	1.141	0.481	(0.319, 2.000)	1.140	1.590	0.913	(0.027, 3.362)
23	1.075	0.490	(0.261, 2.000)	1.005	1.445	0.874	(0.000, 3.119)
24	0.934	0.482	(0.100, 1.872)	0.736	1.128	0.724	(0.000, 2.529)
25	0.931	0.491	(0.070, 1.861)	0.738	1.145	0.756	(0.000, 2.609)

Note: Note that the mode is the same for both priors.

To check the methodology and the `radir` performances, doses are inferred from test data shown in [14] (table 2).

Therefore, the input parameters for the `radir` package should be

```
f <- expression(b0*exp(-b1*exp(-b2*x)) * (1+b3*x*(2*b0*exp(-b1*exp(-b2*x))+1) / (1+b3*x*(b0^2*(exp(-b1*exp(-b2*x)))^2+b0*exp(-b1*exp(-b2*x))))))
pars <- c("b0", "b1", "b2", "b3")
beta <- c(8.6759674, 7.2624173, 0.2296528, 2.3875238)
cov <- matrix(c(0.0562628690, 0.0056047214, -8.120599e-04,
0.018587644, 0.0056047214,
0.0894182387, 9.727568e-04, 0.304724328, -0.0008120599,
0.0009727568, 3.792577e-05,
0.002753902, 0.0185876441, 0.3047243281, 2.753902e-03, 1.14572469
7), nrow=4)
```

Table 2. Cells analyzed and total dicentric counts for the simulated whole body irradiations for testing, and the statistics summary of their respective calibrative densities.

Dose (Gy)	Dicentrics	Cells	Mode	Expected	SD	95% HPD
2	155	498	1.748	1.757	0.110	(1.544, 1.975)
6	425	150	5.850	5.859	0.228	(5.415, 6.308)
12	869	150	9.747	9.763	0.311	(9.158, 10.378)
17	914	100	16.143	16.814	1.778	(14.148, 19.953)

Then, the four situations proposed in table 2 can be introduced in R by means of

```
ex3.a <- dose.distr(f, pars, beta, cov, cells=498, dics=155)
ex3.b <- dose.distr(f, pars, beta, cov, cells=150, dics=425)
ex3.c <- dose.distr(f, pars, beta, cov, cells=150, dics=869)
ex3.d <- dose.distr(f, pars, beta, cov, cells=100, dics=914)
```

Again, the `summary()` function can be used to check that the results are similar to the expected; for instance, for the first experiment we have

```
summary(ex3.a)
Mode
-----
1.748
Expected value
-----
1.757
Standard Dev.
-----
0.11
95% CI
-----
(1.544; 1.975)
```

The probability of a dose exposure between 1.544 and 1.977 is, as expected, approximately 0.95; it can be checked using

```
pr.dose.radir(ex3.a, as.numeric(substr(summary(ex3.a)[[4]],
2, 6)),
as.numeric(substr(summary(ex3.a)[[4]], 9, 13)))
[1] 0.9499784
```

The cumulative distribution for this particular example (figure 4) can be plotted with

```
plot(ex3.a, distr=T)
```

The region under the calibrative curve for doses from 15 to 20 Gy can be plotted for the 17 Gy test data calibrative density (figure 5) using

```
plot(ex3.d, prob=c(15, 20))
```

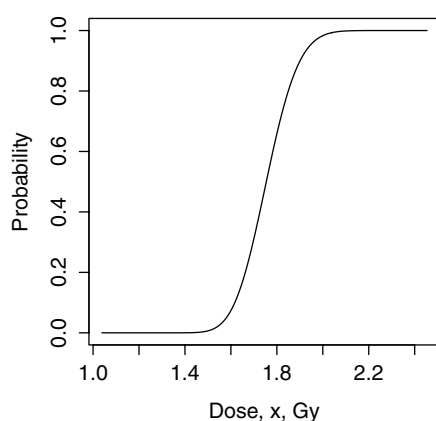


Figure 4. Cumulative distribution function of the 2 Gy test data.

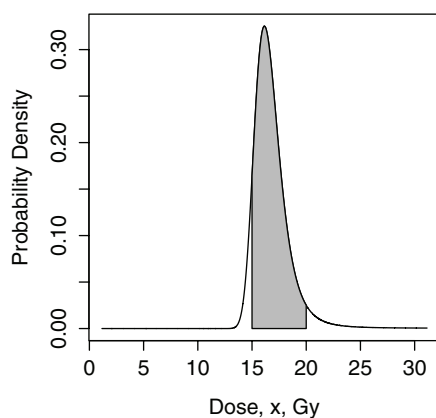


Figure 5. Calibrative density of the 17 Gy test data and the probability of the dose to be in (15, 20).

4. Discussion

Biological dosimetry is necessary in many situations when dealing with radiation events, and a quick and accurate estimation of the radiation doses received by individuals undergoing medical radiation treatments or following a radiation accident is essential in many scenarios. This work presents a readily available package in the framework of the well-known and widely distributed R software that represents a novel and useful tool to achieve this goal. It allows the user to estimate radiation doses received by an individual on the basis of a new inverse regression methodology and using the recently validated Bayesian framework which is able to compute true probability intervals.

The package uses as inputs the estimated parameters and variance–covariance matrix of the dose–response function, obtained using classical (frequentist) maximum likelihood methods. Therefore, the `radir` package can be seen as a complement of other existing packages (CABAS [4], DoseEstimate [5] or BioDoser [6]), which have to be used to obtain the required inputs.

The usability of the `radir` package in several common radiation-related situations has been demonstrated through the proposed examples, although the methodologies introduced in [8] and therefore the package itself could also be used in areas not related to biological dosimetry. Improvements planned for the package `radir` include the consideration of a normal prior dose distribution truncated at negative values [3], the ability to fit the dose-response curve given the calibration data, and the analysis of high-linear energy transform scenarios using the compound Poisson models described in [8]. A Bayesian dose estimation of partial body-irradiated blood samples following the new methodology of [15] will also be considered for future enhancement of the package.

Acknowledgments

This work was funded by the National Institute for Health Research, the grant MTM2012-31118 from the Spanish Ministry of Economy and Competitiveness, and by the grant UNAB10-4E-378 from ERDF ‘A way of making Europe’. The authors are very grateful to the referees and editors for their suggestions, which notably improved the paper.

References

- [1] IAEA 2011 *Cytogenetic Dosimetry: Applications in Preparedness for and Response to Radiation Emergencies* (Vienna: IAEA)
- [2] Ainsbury E A, Vinnikov V A, Puig P, Higuera M, Maznyk N A, Lloyd D C and Rothkamm K 2014 Review of Bayesian statistical analysis methods for cytogenetic radiation biodosimetry, with a practical example *Radiat. Prot. Dosim.* **162** 185–96
- [3] Brame R S and Groer P G 2003 Bayesian methods for chromosome dosimetry following a criticality accident *Radiat. Prot. Dosim.* **104** 61–3
- [4] Deperas J et al 2007 CABAS: a freely available PC program for fitting calibration curves in chromosome aberration dosimetry *Radiat. Prot. Dosim.* **124** 115–23
- [5] Ainsbury E A and Lloyd D C 2010 Dose estimation software for radiation biodosimetry *Health Phys.* **98** 290–5
- [6] Wang H, Liu Q, Wan D, Xiang J, Du L, Wang Y, Cao J, Fu Y, Fan F and Hecker M 2012 BioDoser: improved dose-estimation software for biological radiation dosimetry *Comput. Methods Programs Biomed.* **108** 402–6
- [7] Ainsbury E A, Vinnikov V, Puig P, Maznyk N, Rothkamm K and Lloyd D C 2013 CytoBayesJ: software tools for Bayesian analysis of cytogenetic radiation dosimetry data *Mutation Res.* **756** 184–91
- [8] Higuera M, Puig P, Ainsbury E A and Rothkamm K 2015 A new inverse regression model applied to radiation biodosimetry *Proc. R. Soc.* **471** 20140588
- [9] R Development Core Team 2014 *R: a language and environment for statistical computing* R Foundation for Statistical Computing (Vienna, Austria)
- [10] Morfiña D, Higuera M, Puig P and Oliveira M 2015 The R package `hermite` submitted
- [11] Groer P G and Pereira C A B D 1987 Calibration of a radiation detector: chromosome dosimetry for neutrons *Probability and Bayesian Statistics* ed R Viertl (Berlin: Springer) pp 225–32
- [12] Serna A, Alcaraz M, Navarro J L, Acevedo C, Vicente V and Canteras M 2008 Biological dosimetry and Bayesian analysis of chromosomal damage in thyroid cancer patients *Radiat. Prot. Dosim.* **129** 372–80
- [13] de Keizer B, Hoekstra A, Konijnenberg M W, de Vos F, Lambert B, van Rijk P P, Lips C J M and de Klerk J M H 2004 Bone marrow dosimetry and safety of high ¹³¹I activities given after recombinant human thyroid-stimulating hormone to treat metastatic differentiated thyroid cancer *J. Nucl. Med.: Official Publ., Soc. Nucl. Med.* **45** 1549–54
- [14] Pujol M, Barquinero J F, Puig P, Puig R, Caballín M R and Barrios L 2015 A new model of biodosimetry to integrate low and high doses *PLoS ONE* **9** e114137
- [15] Higuera M, Puig P, Ainsbury E A, Vinnikov V A and Rothkamm K 2015 A new Bayesian model applied to cytogenetic partial body irradiation estimation *Radiat. Prot. Dosim.* (doi:10.1093/rpd/ncv356)