Biodose Tools: An R shiny Application for Biological Dosimetry

Article *in* International Journal of Radiation Biology · February 2023

DOI: 10.1080/09553002.2023.2176564

CITATION
1

READS
104

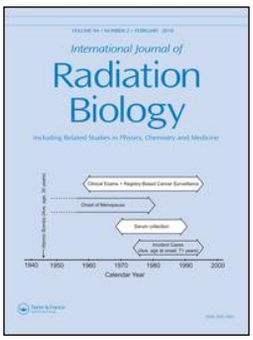
11 authors, including:

David Endesfelder
Bundesamt für Strahlenschutz, BfS München
67 PUBLICATIONS **9,970** CITATIONS
SEE PROFILE

Pedro Puig
Autonomous University of Barcelona
117 PUBLICATIONS **1,629** CITATIONS
SEE PROFILE

Mohamed Amine Benadjaoud
Institut de Radioprotection et de Sûreté Nucléaire (IRSN)
106 PUBLICATIONS **695** CITATIONS
SEE PROFILE

Manuel Higueras
Universidad de La Rioja (Spain)
56 PUBLICATIONS **296** CITATIONS
SEE PROFILE



International Journal of Radiation Biology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/irab20

Biodose Tools: An R shiny Application for Biological Dosimetry

Alfredo Hernández, David Endesfelder, Jochen Einbeck, Pedro Puig, Mohamed Amine Benadjaoud, Manuel Higueras, Elizabeth Ainsbury, Gaëtan Gruel, Ursula Oestreicher, Leonardo Barrios & Joan Francesc Barquinero

To cite this article: Alfredo Hernández, David Endesfelder, Jochen Einbeck, Pedro Puig, Mohamed Amine Benadjaoud, Manuel Higueras, Elizabeth Ainsbury, Gaëtan Gruel, Ursula Oestreicher, Leonardo Barrios & Joan Francesc Barquinero (2023): Biodose Tools: An R shiny Application for Biological Dosimetry, International Journal of Radiation Biology, DOI: 10.1080/09553002.2023.2176564

To link to this article: https://doi.org/10.1080/09553002.2023.2176564

Title: Biodose Tools: An R shiny Application for Biological Dosimetry

Authors: Alfredo Hernández[a], David Endesfelder[b], Jochen Einbeck[c], Pedro Puig[d,e], Mohamed Amine Benadjaoud[f], Manuel Higueras[g], Elizabeth Ainsbury[h], Gaëtan Gruel[i], Ursula Oestreicher[b], Leonardo Barrios[j], Joan Francesc Barquinero[a*]

Affiliations:

[a] Department of Animal Biology, Plant Biology and Ecology (BABVE), Universitat Autònoma de Barcelona, Bellaterra, Spain.

[b] Department of Effects and Risks of Ionising and Non-Ionising Radiation, Federal Office for Radiation Protection, Neuherberg, Germany.

[c] Department of Mathematical Sciences, and Durham Research Methods Centre, Durham University, Durham, UK.

[d] Department of Mathematics, Universitat Autònoma de Barcelona, Bellaterra, Spain.

[e] Centre de Recerca Matemàtica, Bellaterra, Spain.

[f] Radiobiology and Regenerative Medicine Research Service (SERAMED), Institut de Radioprotection et de Sûreté Nucléaire,Fontenay-aux-Roses, France.

[g] Scientific Computation & Technological Innovation Center (SCoTIC), Universidad de La Rioja, Logroño, Spain.

[h] Radiation Effects Department, UK Health Security Agency, Oxford, UK.

[i] Radiobiology of Accidental Exposure Laboratory (LRAcc), Institut de Radioprotection et de Sûreté Nucléaire, Fontenay-aux-Roses, France.

[j] Department of Cell Biology, Physiology and Immunology (BCFI), Universitat Autònoma de Barcelona, Bellaterra, Spain

*Corresponding author: Francesc.Barquinero@uab.cat

Unitat d'Antropologia Biològica, Departament de Biologia Animal, Biologia Vegetal i Ecologia, Universitat Autònoma de Barcelona, E-08193, Bellaterra, Catalonia, Spain

Alfredo Hernández 0000-0002-2660-4545
David Endesfelder
Jochen Einbeck 0000-0002-9457-2020
Pedro Puig 0000-0002-6607-9642
Mohamed Amine Benadjaoud
Manuel Higueras 0000-0001-5530-1714
Elizabeth Ainsbury 0000-0001-8076-6978
Gaëtan Gruel 0000-0001-7856-818X
Ursula Oestreicher
Leonardo Barrios 0000-0002-6151-8503
Joan Francesc Barquinero 0000-0003-0084-5268

Biographical note for each author

Alfredo Hernández, MSc, Physicist and Data Scientist, Independent Researcher, Norwich, UK.

David Endesfelder, Ph.D., Biomathematician, Federal Office forRadiation Protection (BfS), Unit: Biological Dosimetry,Oberschleissheim, Germany

Jochen Einbeck, Ph.D., Associate Professor for Statistics in the Department of Mathematical Sciences, Durham University and Co-Director (Health Data Science) in the Durham Research Methods Centre, Durham, UK.

Pedro Puig, Ph.D., Mathematician, Professor at the Department of Mathematics and member of the Centre de Recerca Matemàtica, Faculty of Sciences, Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallés), Catalonia, Spain.

Mohamed Amine Benadjaoud, Ph.D,. Biomathematician at the Institute for Radiological Protection and Nuclear Safety (IRSN), Radiobiology of Accidental Exposure Laboratory (LRAcc), Fontenay aux Roses, France.

Manuel Higueras, Ph.D., is an associate professor within the Mathematics and Computation Department at the La Rioja University. La Rioja, Spain.

Elizabeth A. Ainsbury, Ph.D., MInstP MSRP CSci, Principal Radiation Protection Scientist and Cytogenetics Group Leader, UK Health Security (UKHSA), UK.

Gaetan Gruel, Ph.D., Researcher and Head of the Laboratory, Institute for Radiological Protection and Nuclear Safety (IRSN), Radiobiology of Accidental Exposure Laboratory (LRAcc), Fontenay aux Roses, France.

Ursula Oestreicher, Dipl. Biol., Biologist and head of the Unit:Biological Dosimetry, Federal Office for Radiation Protection (BfS),Unit: Biological Dosimetry, Oberschleissheim, Germany

Leonardo Barrios, , Ph.D., Biologist, currently retired , former Professor at the Department of Cell Biology, Physiology and Immunology, Faculty of Biosciences, Universitat Autònoma de Barcelona (UAB), Bellaterra (Cerdanyola del Vallès), Catalonia, Spain.

Joan-Francesc Barquinero, Ph.D., Biologist, Associate Professor at the Department of Animal Biology, Plant Biology and Ecology, Faculty of Biosciences, Universitat Autònoma de Barcelona (UAB), Bellaterra (Cerdanyola del Vallès), Catalonia, Spain.

ABSTRACT

Introduction

In the event of a radiological accident or incident, the aim of biological dosimetry is to convert the yield of a specific biomarker of exposure to ionizing radiation into an absorbed dose. Since the 1980s, various tools have been used to deal with the statistical procedures needed for biological dosimetry, and in general those who made several calculations for different biomarkers were based on closed source software. Here we present a new open source program, Biodose Tools, that has been developed under the umbrella of RENEB (Running the European Network of Biological and retrospective Physical dosimetry).

Material and Methods

The application has been developed using the R programming language and the shiny package as a framework to create a user-friendly online solution. Since no unique method exists for the different mathematical processes, several meetings and periodic correspondence were held in order to reach a consensus on the solutions to be implemented.

Results

The current version 3.6.1 supports dose-effect fitting for dicentric and translocation assay. For dose estimation Biodose Tools implements those methods indicated in international guidelines and a specific method to assess heterogeneous exposures. The app can include information on the irradiation conditions to generate the calibration curve. Also, in the dose estimate, information about the accident can be included as well as the explanation of the results obtained. Because the app allows generating a report in various formats, it allows

traceability of each biological dosimetry study carried out. The app has been used globally in different exercises and training, which has made it possible to find errors and improve the app itself. There are some features that still need consensus, such as curve fitting and dose estimation using micronucleus analysis. It is also planned to include a package dedicated to interlaboratory comparisons and the incorporation of Bayesian methods for dose estimation.

Conclusion

Biodose Tools provides an open-source solution for biological dosimetry laboratories. The consensus reached helps to harmonize the way in which uncertainties are calculated. In addition, because each laboratory can download and customize the app's source code, it offers a platform to integrate new features.

KEYWORDS

Biodose Tools; R; biodosimetry; dose assessment; data analysis; graphical user interface

1. Introduction

Biological dosimetry aims at estimating the absorbed dose in an individual in which an exposure to ionizing radiation (IR) is suspected, by means of analyzing biomarkers with a clear dose-effect relationship (IAEA 2011). A great majority of biomarkers of dose exposure come from the analysis of the induced DNA damage, most of them analyzed using cytogenetic techniques such as dicentric, translocation, or micronucleus assays. Dose assessment is based on converting an observed yield of aberrations (e.g., the frequency of dicentrics present in peripheral blood lymphocytes) into an absorbed dose using a pre-established calibration curve. This process requires mathematical models and the assumptions on statistical probability distribution of the aberration in question. First, to establish a calibration curve, blood samples must be uniformly irradiated at several doses and the observed distribution of aberrations is mostly assumed to follow a Poisson distribution. More precisely, it is assumed that for low LET (linear energy transfer) radiation types, uniform exposures result in dicentric counts that follow a Poisson distribution (Bauchinger and Schmid 1973; Edwards et al. 1979). High LET radiation types tend to show overdispersion (Virsik and Harder 1981; Brame and Groer 2002; Puig and Barquinero 2011). Overdispersion also often arises in micronuclei counts and γ-H2AX foci data (Vral et al. 2011; Einbeck et al. 2018) and for semi-automaticaly socted dicentrics (Endesfelder et al. 2020) . Then the observed yields at different doses are used to construct a calibration curve assuming a Linear-Quadratic (LQ) or Linear (L) model, depending on the radiation quality. While low-LET exposures show a linear-quadratic dose-effect relationship, high-LET exposures tend to be linear, (Edwards et al. 1980). The coefficients of these models can be estimated using maximum likelihood or iteratively reweighted least squares approaches (Merkle 1983).

In case of a radiation accident, the observed distribution of aberration counts is tested for deviations from the Poisson distribution to distinguish between homogeneous and heterogeneous exposures. The observed count is inverse regressed by the calibration curve and uncertainties are usually calculated considering only the uncertainty relative to the yield observed or considering also uncertainties coming from the calibration curve (Edwards 1978; Merkle 1983; Savage et al. 2000). Procedures to consider both sources of uncertainties are not simple, as the error is made up of two components: (a) uncertainties from the distribution of observed counts, and (b) uncertainties from the LQ or L models of the calibration curve (Merkle 1983; Savage et al. 2000). In case of a heterogeneous exposure, additional distributional assumptions have to be made (Sasaki and Miyata 1968; Dolphin 1969; Pujol et al. 2016). Bayesian methods have been recently proposed to consider both sources of uncertainty, and for both, whole-body (homogeneous) and partial-body (heterogeneous) dose assessment (Ainsbury et al. 2014; Higueras et al. 2015; Moriña et al. 2015).

1.1. *Software implementations*

The tools to deal with some or all of the statistical procedures used in biological dosimetry have evolved since the 1980s (see Ainsbury and Barquinero 2009) . These tools ranged from adapted Microsoft Excel sheets to specific programs (the later summarized in Table 1). The specific programs were based on closed-source software and, for example, used a single methodology to calculate uncertainties. Thus, improvements were restricted to those who manage the source code, and the end-user cannot modify or implement any improvement. In addition, there is a possibility of misunderstanding or misuse of these closed programs.

Insert table 1

Recently, using the R project for statistical computing (R Core Team 2022) some scripts have been written for biological dosimetry purposes. Although R programming is based on open-source code, most end-users performing biodosimetry are usually not familiar with programming and mathematical or statistical modeling that is required to use or implement R source code. It is therefore required to provide tools with GUIs that are easy to use and that provide the required functionality to obtain reliable dose estimates.

In this paper we present biodosetools (styled as Biodose Tools), an R package developed using shiny (Chang et al. 2022), to offer an online and easy-to-use solution to be used by biological dosimetry laboratories, as well as a tool for statisticians to manually perform the underlying calculations in R. The package is available from the Comprehensive R Archive Network (CRAN) at https://cran.r-project.org/web/packages/biodosetools/index.html, and GitHub at https://github.com/biodosetools-team/biodosetools and can be distributed under the GPL-3.0 license (GNU General Public License v3.0). Biodose Tools has been developed under the umbrella of RENEB (Running the European Network of Biological and retrospective Physical dosimetry).

2. Materials and methods

2.1. *Statistical considerations*

In biological dosimetry, the quantity being modeled is usually the *yield*, i.e. the mean aberration count per cell. For low-LET radiation, the yield (λ) of chromosome aberrations (e.g. dicentrics, translocations, micronuclei) is related to exposed dose (D) by the LQ equation:

$$\lambda = C + \alpha D + \beta D^2 \qquad (1)$$

For high-LET radiation, as well as for highly protracted low-LET exposures (see Section 2.2.4.), the α-term becomes large, and eventually, the β-term becomes biologically less relevant and also statistically "not significant" (Frome and DuFrain 1986; Sasaki 2003). In this situation, the dose-effect is approximated by the linear equation (2):

$$\lambda = C + \alpha D \qquad (2)$$

In both equations, C is interpreted as the background level.

2.1.1. Dose-effect curve fitting

The objective of dose-effect or curve fitting is to calculate the coefficients C, α and β which best fit the calibration data points. For dicentrics (for more details see Section 2.2), simulated whole-body irradiation of ex vivo blood in test tubes with X-rays or γ-rays produces aberration counts which are very well represented by the Poisson distribution (Edwards et al. 1979). In contrast, neutrons and other types of high-LET radiation produce distributions which display overdispersion, where the sample variance exceeds the sample mean. Rather

than choosing a more suitable distribution, this is usually dealt with by adjusting the calculating the overdispersion and then adjusting the uncertainty accordingly (IAEA 2011).

2.1.2. Testing for Poisson

Because most curve fitting methods rely on Poisson statistics (see Supplementary Materials Section 1.1), the dicentric cell distribution should be checked for Poisson compliance for each dose used to build the calibration curve. This should also be checked for the sample tested for an exposure. Although recently in the field of biological dosimetry different tests have been proposed to check the Poisson distribution (Duran et al. 2002; Higueras et al. 2018), the most widely used test is the *u*-test (Rao and Chakravarti 1956; Savage 1970). The *u*-test statistic (3) is a normalized unit of the dispersion index $(\hat{\sigma}^2/\bar{y})$, which for a Poisson distribution should be close to 1:

$$u = (\hat{\sigma}^2/\bar{y} - 1)\sqrt{\frac{N-1}{2(1-1/X)}} \quad (3)$$

When $-1.96 \leq u \leq 1.96$ the assumption of equidispersion is not rejected with a two-tailed significance level of $\alpha = 0.05$. A *u*-value higher than 1.96 indicates overdispersion, whilst *u*-values lower than -1.96 indicate underdispersion. Underdispersion is an extraordinary occurrence in Biodosimetry. However, Pujol et al. (2014) claimed that, after high dose exposure (> 5 Gy), the underdispersion detected in their datasets was caused probably for two reasons. The first, due to the number of chromosomes that the cells possess and the limitation of forming cells with a very high number of dicentrics. The second is that at high doses fewer

cells with 0 or 1 dicentric than expected were observed, possibly indicating a lower efficiency in repairing genetic damage at high doses.

2.2. *Dose-effect curve fitting*

2.2.1. Fitting method

Maximum likelihood is the method proposed for determining the best fit (Papworth 1975; Merkle 1983). Using this method, the best fit value for each coefficient is achieved by assuming a Poisson distribution and maximizing the likelihood of the observations by the method of iteratively reweighted least squares. In case of overdispersed distributions, as obtained after high LET radiation, the weights must take into account the overdispersion. Table 2 gives example data used to construct a dose-effect curve for low LET γ-radiation (Barquinero et al. 1995), while Table 3 shows the estimated coefficients when the data from Table 2 are fitted. The p-values of the t-test indicate that each parameter is statistically significant (see Supplementary Materials Section 1.2 for more details on testing goodness of fit).

Insert table 2

Insert table 3

2.2.2. Dose-effect calibration

Adequate curve fitting requires a sufficient number of evaluated dose points to minimize the error. A minimum of 7 doses should be evaluated, 5 of them at doses equal or lower than 1 Gy, including the 0

Gy dose. Usually, for low LET radiation (X-rays and γ-rays) the dose range evaluated is 0–5 Gy, at higher doses, evidence of saturation of the aberration yield exists. This saturation will result to a distortion of the β coefficient in equation (1) (Lloyd and Edwards 1983; ISO019328:2014). For high LET radiation, a range of 0–2 Gy is suggested (IAEA 2011). At higher doses, scoring should aim to detect 100 dicentrics at each dose. However, at lower doses this is difficult to achieve and instead several thousand cells per point should be scored; a number between 3000 and 5000 is suggested. In all cases, the actual number of cells scored should be dependent on the number of dose points in the low dose region, with the focus on minimizing the error on the fitted curve (Higueras et al. 2020).

Opinions vary on how to deal with the background level of chromosomal aberrations when fitting dose-effect data. In general, there are three approaches: (a) a dose point at 0 Gy is included in the curve fitting procedure, (b) the zero dose point is ignored, or else (c) the zero dose point is represented in every fitting procedure by a standard background value. Nowadays almost all laboratories include a 0 Gy dose point, and it should always be included for dicentrics, translocations and micronuclei. If the measured yield at zero dose is used as one of the data points for the curve fitting (as used in the curve fitting presented in Table 3), the background becomes a variable parameter. However, since the yield in unirradiated cells is usually low, often none are observed so the measured yield at zero dose is zero. As discussed, at low doses, the statistical resolution of the data points is generally low. Thus, including the zero dose point in the curve fitting procedure can in some cases lead to negative estimated C and α coefficients, which obviously have no biological basis. A way to resolve this problem is to ignore the zero dose data point and constraining the curve to pass through the origin. Another way to solve this is to fit the calibration curve model with constrained maximum likelihood which forces the intercept parameter to be non-negative (Oliveira et al. 2016). Some experts have used a small positive background value as a data point and assigned a large percentage of uncertainty to it. Ideally each laboratory should generate its own background data. A consensus has emerged that the background

level of dicentrics is ~ 0.5–1.0 per 1000 cells (Lloyd et al. 1980) whilst for translocations (Sigurdson et al. 2008) and for micronuclei (Fenech 1993) the control values are higher.

2.2.3. Protracted and fractionated exposure

In case of a protracted or fractionated exposure to low-LET radiation the resulting chromosomal aberration yield may be lower than when receiving the same dose acutely (Edwards et al. 1980). Dose protraction reduces the dose squared coefficient, β, in equation (1). It is assumed that this term refers to those aberrations generated by the interaction of two or more tracks. Exchange type aberrations, such as dicentrics have a fast kinetics formation (Darroudi et al. 1998; Durante et al. 1996, 1998; Pujol-Canadell et al. 2020), and hence the effect on the interaction of two tracks can be modified by repair mechanisms that have time to function during chronic or intermittent acute exposures. In the early times of radiobiology and evaluating the effect of IR on *Tradescantia* Lea and Catcheside (1942) suggested a time-dependent factor called the *G* function (4) to allow change of the dose squared coefficient and so account for the impacts of dose protraction.

$$G(x) = \frac{2}{x^2}(x - 1 + e^{-x}), \quad \text{and } x = \frac{t}{t_0} \qquad (4)$$

where t is the time over which the irradiation occurred, and t_0 is the mean lifetime of the breaks, which has been shown to be of the order of ~ 2 hours (Bauchinger et al. 1979; Lloyd et al. 1984).

Currently in biological dosimetry this approach is still accepted and that is why the LQ dose-effect equation (1) may be modified as follows:

$$\lambda = C + \alpha D + \beta G(x) D^2 \qquad (5)$$

Given a value of *t*, we can have three protraction cases:

a) Acute exposure when *t* is approximately 0. In this case $\lim_{x \to 0} G(x) = 1$. Therefore, the dose-effect equation becomes (1).

b) Highly protracted exposure when *t* is high. In this case *G*(*x*) reduces virtually to zero. Therefore, even if a high dose (> 1.0 Gy) is involved, the dose-effect equation becomes linear (2).

c) Protracted exposure for any case in between. In this case the effects of *G*(*x*) need to be considered.

All dose estimation methods implemented in biodosetools consider protraction, with the extreme of *G*(*x*) = 1 or *G*(*x*) ≈ 0 being just particular cases. For this reason, all methods described in Section 2.3 are derived using equation (5) instead of (1).

2.2.4. Genomic equivalence for translocations

Curve fitting for translocations follow the same procedure as described for dicentrics. However, Fluorescence *In Situ* Hybridization (FISH) analysis only evaluates the translocation frequency in a specific set of chromosomes painted. The conversion of this frequency to full genome equivalence is a recommended procedure to use when different data must be combined or when results using different combinations of whole chromosome paintings are compared, e.g. in the frame of interlaboratory comparisons.

When the DNA probes used to paint different chromosome pairs are labelled with different fluorochromes, the genomic conversion factor F_p/F_G is usually calculated by using the formula (6) for the painted fractions of the genome (Lucas and Deng 2000):

$$\frac{F_p}{F_G} = \frac{2}{0.974}\left[\sum_i f_i(1-f_i) - \sum_{i<j} f_i f_j\right] \quad (6)$$

where F_G is the full genome aberration frequency, F_p is the translocation frequency detected by FISH, and f_i is the fraction of genome corresponding to each chromosome i used in the hybridization, taking into account the gender of the subjects (Morton 1991). The total number of interchromosomal exchanges is 0.974, using the same assumption of DNA proportionality (see calculations in Lucas et al. 1992).

Translocations have higher background levels than dicentrics, which is partly due to the former being a persistent type of aberration. When attempting retrospective biological dosimetry, it is critical to account for the translocation background, especially at low doses. Because a pre-exposure control blood sample from the unintentionally irradiated person or a population study group is not available, an estimated value based on generic survey data must be utilized. Ideally, a laboratory would create its own control database, but this is a huge undertaking because it would have to cover a lot of confounding factors and, more importantly, a wide range of age groups. The greatest international database, split down by age, sex, ethnicity, geographic region, and smoking behaviors, is currently available thanks to a comprehensive meta-analysis published by Sigurdson et al. (2008). It is important to

account for the background and subtract from the total number of translocations observed in an individual's lymphocyte the expected translocation rate given a set of confounding factors (7), the most important of which is age:

$$F_G' = \frac{X - X_c}{N(F_p/F_G)} \quad (7)$$

where F_G' is the corrected full genome aberration frequency, X is the total number of translocations observed, X_c is the total number of expected aberrations, N is the total number of cells, and F_p/F_G is the genomic conversion factor given by (6).

2.3. *Uncertainty on dose estimation*

While it is simple to determine a dose from a measured yield of dicentrics, the associated uncertainty can be calculated using a variety of methods. Calculating 95% confidence limits is a common procedure for expressing uncertainty in terms of a confidence interval. A frequentist 95% confidence interval means that with a large number of repeated samples, 95% of such calculated confidence intervals would include the true value of the dose. The problem in estimating confidence limits for dicentrics and translocations after low LET exposure comes from two sources of uncertainty: the uncertainty from the aberration yield of the person to be examined, and uncertainties associated with the calibration curve. This issue has been discussed in the literature (Merkle 1983; Savage et al. 2000; Szłuińska et al. 2007).

Depending on the type of exposure, methods for whole-body assessment, partial-body

assessment, and assessments for heterogeneous exposures with two different doses were implemented in the biodosetools package, using the following methods. The biodosetools package implements the following assessment methods, which are indicated in IAEA 2001 and IAEA 2011:

- Whole-body assessment: Merkle's method (Merkle 1983).
- Whole-body assessment: delta method (IAEA 2001).
- Partial-body assessment: Dolphin's method (Dolphin 1969).

In addition biodosetools also includes a package to assess heterogeneous exposures using a mixed Poisson model (Pujol et al. 2016). That was used in a recent RENEB/EURADOS field exercise (Endesfelder et al. 2021). In-depth statistical descriptions of each method is available in Supplementary Materials Section 1.3.

3. Results

3.1. *Package implementation*

The development of biodosetools started in 2018 in the frame of RENEB (Running the European Network of Biological and retrospective Physical dosimetry), one of whose purposes is to standardize methodology for individualized dose estimation from biological methods to statistical ones. Since no unique method exists for the different mathematical processes involved in biological dosimetry, several meetings and periodic correspondence were held in order to reach a consensus on the solutions to be implemented and, in particular, the solutions to be proposed as default.

As of version 3.6.1 (November 2022), biodosetools supports dose-effect fitting and dose estimation for the following biodosimetry assays:

- Dicentric assay.
- Translocation assay.

The biodosetools package is available from CRAN https://CRAN.R-project.org/package=biodosetools and can be installed using the standard R tools. In Table 4 we list all the functions available to the user through the R API provided by biodosetools.

Insert table 4

3.1.1. Used packages

Naturally, biodosetools is built on other packages. Data transformations and wrangling are done through dplyr (Wickham et al. 2022), tidyr (Wickham and Girlich 2022), rlang (Henry and Wickham 2022), and magrittr (Bache and Wickham 2022), while visualizations are done using ggplot2 (Wickham 2016). All of the aforementioned packages are part of the tidyverse meta-package (Wickham et al. 2019). While most statistical calculations are implemented ad-hoc with the stats (R Core Team 2022) package, we rely in additional packages for some statistical analyses, such as maxLik (Henningsen and Toomet 2011), mixtools (Benaglia et al. 2009), MASS (Venables and Ripley 2002), and msm (Jackson 2011). For improved legibility, messages and warnings in the command line interface are constructed using cli (Csárdi 2022).

3.1.2. User interface using shiny

The Biodose Tools user interface is written in R shiny (Chang et al. 2022) using Bootstrap 3, via the shinydashboard (Chang and Borges Ribeiro 2021), shinyWidgets (Perrier et al. 2022), and bsplus (Lyttle 2021) packages, and the golem (Fay et al. 2022) framework. Analyses are implemented in the R programming language (R Core Team 2022), with the resultant tables and plots rendered in HTML through JavaScript libraries, via either native shiny or rhandsontable (Owen 2021). This is done by the JavaScript engine shipped on browser of choice, or by an instance of QtWebKit if the app is run within RStudio. The shiny app allows to download processed count or case data via xtable (Dahl et al. 2019). Similarly, it allows to generate and download reports, which are rendered using rmarkdown (Allaire et al. 2022) and Pandoc (MacFarlane 2022) to convert a native .Rmd document directly into an Office Open XML .docx document, or a .pdf file built from an intermediary .tex file using LaTeX (Lamport 1994).

In Biodose Tools color is used to identify different sections of information, a technique called color-coding. In particular, Biodose Tools uses four colors to code information: (a) blue for options and settings, (b) purple for data input, (c) green for results, and (d) orange for exporting results. Figure 1 shows a simple mock-up of how the four types of boxes and corresponding widgets would fit together in one of the shiny app's modules.

Insert figure 1

3.1.3. Running Biodose Tools

For an offline use of Biodose Tools, we need to run the following commands in R console:

```
R> library(biodosetools)
R> run_app()
```

A series of video tutorials have been prepared with the aim of helping biodosetools users in the installation of R, RStudio, and the package itself, as well as in the usage of the shiny user interface to perform dose-effect fitting and dose estimation. They can be found on our YouTube channel: https://www.youtube.com/@biodosetools. The "Installation Guide" tutorial has captions available in 8 languages. And as of 14th November, 2022 the number of views of the three existing tutorials has been 508.

The following examples illustrate the functionality of biodosetools's shiny user inter- face to perform dose-effect fitting and dose estimation for the dicentric assay. The equivalent examples using the R API can be found in Supplementary Materials 2. Additional examples are available on https://biodosetools-team.github.io/documentation/.

3.2. *Dicentrics dose-effect fitting*

3.2.1. Input count data

The first thing to do is enter the count data. Biodose Tools allows to enter the count data from a file or manually (supported file formats are .csv, .dat, and .txt) (Figure 2). Once the table is generated and filled (Figure 3), the "Calculate parameters" button will calculate the total number of cells (N), total number of aberrations (X), as well as mean ($\bar{y}$) variance (σ^2), dispersion index ($\hat{\sigma}^2/\bar{y}$), and u-value.

Insert figure 2

Insert figure 3

3.2.2. Irradiation conditions

As it is known that the irradiation conditions during calibration influence future dose estimates (Trompier et al. 2017), and to improve their traceability, the user can enter the conditions in which the samples used to build the curve were irradiated. This option is only available in the shiny app (Figure 4), so that these can be saved into the generated reports.

Insert figure 4

3.2.3. Perform fitting

For the calibration adjustment it is necessary to select those options that are the most suitable, for example choosing between L or LQ adjustment (Figure 2), and then click on the "Calculate fitting" button on the "Data input" box (Figure 3). Fitting results and summary statistics will appear in the "Results" tabbed box, and a graphical representation of the dose-effect curve is displayed in the "Curve plot" box (Figure 5).

Insert Figure 5

The "Export results" box (Figure 5) displays two buttons: (a) "Save fitting data", and (b)

"Download report". The "Save fitting data" will generate an .rds file that includes all information about the count data, irradiation conditions, and options selected when performing the fitting. This file can then be loaded in the dose estimation module to load the dose-effect curve coefficients. Similarly, the "Download report" will generate a .pdf or a .docx report containing all inputs and fitting results.

3.3. *Dicentrics dose estimation*

3.3.1. Load pre-calculated curve

The first step is to either load the pre-calculated curve in .rds format obtained in the dose-effect fitting module (Figure 6) or input the curve coefficients manually (Figure 7) in case the user wants to use a pre-existing curve calculated outside of Biodose Tools. Clicking on "Preview data" will load the curve into the app and display it on the "Results" tabbed box.

Insert Figure 6

Insert Figure 7

3.3.2. Input case data

Next, we can choose to enter the count data from a file or manually (the supported file formats are .csv, .dat, and .txt) (Figure 8). Once the table is generated, clicking the "Calculate parameters" button will calculate the total number of cells (N), total number of aberrations (X), as well as mean ($\bar{y}$), standard error (σ), dispersion index ($\hat{\sigma}^2/\bar{y}$), and u-value. The shiny app

also includes the option to include the information about the incident that is being evaluated. This information may be relevant to explain the results obtained and is included in the generated reports.

Insert Figure 8

3.3.3. Perform dose estimation

The final step is to select the dose estimation options depending on the characteristics of the accident. In the "Dose estimation options" box (Figure 9) we can select type of exposure (acute, protracted, and highly protracted), type of assessment (whole-body, partial-body, or heterogeneous), and error methods for each type of assessment.

Insert figure 9

The dose estimation results are shown in Figure 10. Once the estimation is done the app also incorporates the possibility to describe the results obtained in the "Save results" box. All information, the calibration curve used, the data of the case, the different estimated doses, as well as the description of the case and the interpretation of the results can be saved generating a .pdf or a .docx report via the "Download report" button. It is important to note that biodosetools can be used not only to estimate the doses, but also to draft a report of the accident being evaluated with full traceability. This can be further adapted and customized to each laboratory's internal needs thanks to the open-source nature of the project.

Insert figure 10

3.4. *Technical matters*

Biodose Tools has been developed following international guidelines of biological dosimetry (IAEA 2011; ISO19238:14; ISO20046:2019). Despite our best efforts to ensure that biodosetools automatically handles mathematical errors behind the scenes, such as using a constraint maximum-likelihood optimization method (Oliveira et al. 2016) when the fitting using a generalized linear model (GLM) is not possible, or correcting negative dose estimates, the resulting curves may still result in errors (both using the shiny app or the R API) when performing dose estimation. As an example, let us consider the following curve, shown in Table 5 and Figure 11. Although at a glance we can already suspect that there is some kind of issue in the observed counts, due to the low number of evaluated cells, one may still decide to proceed and use it for dose estimations.

Insert Table 5

Insert Figure 11

In Table 6 below we can see an example of a case we may want to estimate the dose for. If we try to estimate the dose using Merkle's method, we encounter an error involving the

uniroot() function. The function uniroot() searches for a root (i.e., zero) of the function $f(x_1, \ldots, x_n)$ with respect to its first argument x_1 within a specified interval. In our example, this error occurs when projecting the upper 95% confidence limit of the yield λ_U into the lower curve (S1.6). The reason for this is that the lower 95% confidence band of the dose-effect curve is not a monotonically increasing function, meaning that D_L has multiple possible numerical solutions, as shown in Figure 12.

Insert table 6

Insert Figure 12

It is worth noting that this projection error is particular to Merkle's method for whole- body assessment and could certainly be circumvented if one chooses to use the delta method instead. However, for optimal dose estimation we expect both the calibration curve and its confidence intervals to define monotonically increasing functions.

4. Discussion

Biological dosimetry has undergone a rapid evolution since the publication of the last IAEA manual (IAEA 2011). On the one hand, the potential of early biomarkers of exposure to ionizing radiation has been evaluated. Such as detection of γ-H2AX foci and gene expression (Kulka et al. 2017; Abend et al. 2021). In addition, new statistical approaches have been proposed: checking the Poisson distribution (Higueras et al., 2018); to estimate dose using Bayesian approaches (Ainsbury et al. 2014; Higueras et al. 2015; González et al. 2020); as well as being able to evaluate laboratory performance in interlaboratory comparison (ILC)

exercises (González et al. 2022). In some cases these improvements have also been translated to shiny apps (https://manu2h.shinyapps.io/InterLabComparison/, https://manu2h.shinyapps.io/gof_poisson/).

The intention of the Biodose Tools team is to expand the capabilities of the app, and for Biodose Tools to serve as a platform to improve mathematical interpretation in biological dosimetry. Currently, there are a few improvements that we are already working on for future biodosetools releases, such as allowing .xlsx and .xls for data input, as well as full support for micronuclei assay analysis. Considering the periodic ILCs that the biological dosimetry community performs, another major future improvement is a specific module to standardize the dose estimation and data handling for ILCs. Last but not least, there are plans to integrate Bayesian dose estimation methods into biodosetools (Higueras et al. 2015), as well as to integrate existing software for γ-H2AX dose estimation (Einbeck et al. 2018) into biodosetools, and to add support for multiple count or case data inputs.

Biodose Tools is a joint international effort between mathematicians, statisticians, and cytogeneticists from different labs, which allowed us to arrive at a consensus on the classical or most commonly used statistical methods to be used in each step of the fitting and dose estimation processes. Due to the package and shiny application being open source and modular, expanding or changing the methods is a trivial matter. The shiny app, and the reports allow complete statistical traceability, which is very important for new laboratories that want to be accredited by the ISO standards: ISO 21243:2008 for triage, ISO 19238:2014 for dicentrics, ISO 20046:2019 for translocations (ISO21243:2008, ISO19238:2014,

ISO20046:2019).

As of now, Biodose Tools is being used globally in different exercises and trainings, such as in the Lund exercise performed by RENEB and EURADOS working group 10 (between November 2019 and March 2020) (Endesfelder et al. 2021) as well as in a IAEA course in Thailand (30th July – 2nd August, 2019), where participants showed very positive feedback. The current RENEB 2021 interlaboratory comparison involving 85 specialized labs from 27 countries was performed exclusively using biodosetools 3.5; presented at RADRES's annual meeting (3rd – 6th October 2021) (Port 2021).

Acknowledgments

We would like to thank all members of the RENEB association who helped in the design and testing of the package and Shiny app. This research was done with support from a grant from the Spanish government (ref. PCIN-2017-077) and a grant from the Consejo de Seguridad Nuclear Español (BOE-A-2019-311).

We would also like to thank everyone outside of the Biodose Tools team who helped providing subtitles for our video tutorials on YouTube, in particular Tomisato Miura from Hirosaki University, who helped with the Japanese subtitles, and Tai-I Chen, who helped with the Chinese (Simplified) and Chinese (Traditional) subtitles.

Disclosure statement

No conflict of interest was reported by the authors.

References

Abend M, Amundson SA, Badie C, Brzoska K, Hargitai R, Kriehuber R, Schüle S, Kis E, Ghandhi SA, Lumniczky K, et al. 2021. Inter-laboratory comparison of gene expression biodosimetry for protracted radiation exposures as part of the RENEB and EURADOS WG10 2019 exercise. Sci. Rep. 11(1):9756.

Ainsbury EA, Barquinero JF. 2009. Biodosimetric tools for a fast triage of people accidentally exposed to ionizing radiation. Statistical and computational aspects. Ann Ist Super Sanita. 45(3): 307–312.

Ainsbury EA, Lloyd DC. 2010. Dose estimation software for radiation biodosimetry. Health Phys. 98(2), 290–295.

Ainsbury EA, Vinnikov VA, Puig P, Higueras M, Maznyk NA, Lloyd DC, Rothkamm K. 2014. Review of bayesian statistical analysis methods for cytogenetic radiation biodosimetry, with a practical example. Radiat Prot Dosimetry. 162(3): 185–196.

Allaire JJ, Xie Y, McPherson J, Luraschi J, Ushey K, Atkins A, Wickham H, Cheng J, Chang W, Iannone R. 2022. rmarkdown: Dynamic Documents for R. https://github.com/rstudio/rmarkdown

Bache S. M, Wickham H. 2022. magrittr: A Forward-Pipe Operator for R. https://magrittr.tidyverse.org.

Barquinero JF, Barrios L, Caballín MR, Miró R, Ribas M, Subias A, Egozcue J. 1995. Establishment and validation of a dose-effect curve for γ-rays by cytogenetic analysis. Mutat Res. 326(1): 65–69.

Bauchinger M, Schmid E. 1973. Chromosome aberrations in human lymphocytes after X-irradiation in vitro II. Analysis of primary processes in the formation of dicentric chromosomes. Mutat Res. 20(1): 107–113.

Bauchinger M, Schmid E, Dresp J. 1979. Calculation of the Dose-rate Dependence of the Dicentric Yield after Co γ-irradiation of Human Lymphocytes. Int J Radiat Biol Relat Stud Phys Chem Med. 35(3): 229–233.

Benaglia T, Chauveau D, Hunter DR, Young D. 2009. mixtools : An R Package for Analyzing Finite Mixture Models. J Stat Softw. 32(6), 1–29.

Brame RS, Groer PG. 2002. Bayesian Analysis of Overdispersed Chromosome Aberration Data with the Negative Binomial Model. Radiat Prot Dosimetry. 102(2):115–119.

Bundesamt für Strahlenschutz. (1996). MLREG: ML-Regression.

Chang W, Borges Ribeiro B. 2021. shinydashboard: Create Dashboards with Shiny. http://rstudio.github.io/shinydashboard/

Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, Allen J, McPherson J, Dipert A, Borges B. 2022. shiny: Web Application Framework for R. https://shiny.rstudio.com/

Csárdi G. 2022. cli: Helpers for Developing Command Line Interfaces. https://cli.r-lib.org

Dahl DB, Scott D, Roosen C, Magnusson A, Swinton J. 2019. xtable: Export Tables to LaTeX or HTML. http://xtable.r-forge.r-project.org/

Darroudi F, Fomina J, Meijers M, Natarajan AT. 1998. Kinetics of the formation of chromosome aberrations in X-irradiated human lymphocytes, using PCC and FISH. Mutat. Res. - Fundam. Mol. Mech. Mutagen. 404(1–2): 55–65.

Deperas J, Szłuińska M, Deperas-Kaminska M, Edwards A, Lloyd D, Lindholm C, Romm H, Roy L, Moss R, Morand J, et al. (2007). CABAS: a freely available PC program for fitting calibration curves in chromosome aberration dosimetry. Radiat Prot Dosimetry. 124(2): 115–123.

Dolphin GW. 1969. Biological Dosimetry with Particular Reference to Chromosome Aberration Analysis: A Review of Methods. In International Atomic Energy Agency (Ed.), Handling of Radiation Accidents (pp. 215–224).

Duran A, Barquinero JF, Caballín MR, Ribas M, Puig P, Egozcue J, Barrios, L. 2002. Suitability of FISH painting techniques for the detection of partial-body irradiations for biological dosimetry. Radiat Res. 157(4), 461–468.

Durante M, Furusawa Y, Gotoh E. 1998. Technical Report A simple method for simultaneous interphase-metaphase chromosome analysis in biodosimetry. Int J Radiat Biol. 74(4): 457–462.

Durante M, George K, Yang TC. 1996. Biological Dosimetry by Interphase Chromosome Painting. Radiat Res., 145(1): 53-60.

Szłuińska M, Edwards A, Lloyd D. 2005 Statistical methods for biological dosimetry HPA-RPD-011 https://www.gov.uk/government/publications/biological-dosimetry-statistical-methods

Edwards AA. 1978. The estimation of dose and its uncertainty. In: NRPB-R 70, doses in radiation accidents investigated by chromosome aberration analysis, VIII A review of cases investigated: 1977. Lloyd DC, Purrott RJ, Prosser JS, Edwards AA, Dolphin GW, White AD, Reeder EJ, White CM,

Copper SJ, Stephenson BD, Tipper PA (eds) National Radiation Protection Board, Didcot, Oxon, pp 16–17

Edwards AA, Lloyd DC, Purrott RJ. 1979. Radiation induced chromosome aberrations and the poisson distribution. Radiat Environ Biophys. 16(2), 89–100.

Edwards AA, Lloyd DC, Purrott RJ. 1980. Dicentric chromosome aberration yield in human lymphocytes and radiation quality; A resume including recent results using alpha-particles. Seventh Symposium on Microdosimetry, Report EUR7147, 1263–1273.

Einbeck J, Ainsbury EA, Sales R, Barnard S, Kaestle F, Higueras M. 2018. A statistical framework for radiation dose estimation with uncertainty quantification from the γ-H2AX assay. PLoS One, 13(11), e0207464.

Endesfelder D, Kulka U, Einbeck J, Oestreicher U. 2020. Improving the accuracy of dose estimates from automatically scored dicentric chromosomes by accounting for chromosome number. Int J Radiat Biol. 96(12): 1571–1584.

Endesfelder D, Oestreicher U, Kulka U, Ainsbury EA, Moquet J, Barnard S, Gregoire E, Martinez JS, Trompier F, Ristic Y, et al. 2021. RENEB/EURADOS Field Exercise 2019: Robust Dose Estimation under outdoor Conditions based on the Dicentric Chromosome Assay. Int J Radiat Biol. 97(9), 1181–1198.

Fay C, Guyader V, Rochette S, Girard C. 2022. golem: A Framework for Robust Shiny Applications. https://github.com/ThinkR-open/golem

Fenech M. 1993. The cytokinesis-block micronucleus technique: A detailed description of the method and its application to genotoxicity studies in human populations. Mutat. Res. - Fundam. Mol. Mech. Mutagen. 285(1): 35–44.

Frome EL, DuFrain RJ. 1986. Maximum Likelihood Estimation for Cytogenetic Dose-Response Curves. Biometrics, 42(1), 73-84.

Garcia L O and Zequera JT. 1996. Mathematical operations in cytogenetic dosimetry: Dosgen. Proceedings of the 3 Regional Meeting on Radiological and Nuclear Safety Radiological protection in Latin America and the Caribbean, (p. 1207).

González JE, Barquinero JF, Holladay BA, di Giorgio M, Higueras M. 2020. Uncertainty calculation methods in dose assessment for dicentric chromosome assay. Int J Radiat Biol. 96(5): 606–613.

González JE, Holladay B, Higueras M, di Giorgio M, Barquinero JF. 2022. Assessment methods for

inter-laboratory comparisons of the dicentric assay. Int J Radiat Biol. Jul 20: 1–8.

Henningsen A, Toomet O. 2011. maxLik: A package for maximum likelihood estimation in R. Comput Stat. 26(3): 443–458.

Henry L, Wickham H. 2022. rlang: Functions for Base Types and Core R and Tidyverse Features. https://rlang.r-lib.org

Higueras M, González JE, di Giorgio M, Barquinero JF. 2018. A note on Poisson goodness-of-fit tests for ionizing radiation induced chromosomal aberration samples. Int J Radiat Biol. 94(7): 656–663.

Higueras M, Howes A, López-Fidalgo J. 2020. Optimal experimental design for cytogenetic dose-response calibration curves. Int J Radiat Biol. 96(7): 894–902.

Higueras M, Puig P, Ainsbury EA, Vinnikov VA, Rothkamm K. 2016. A new Bayesian model applied to cytogenetic partial body irradiation estimation. Radiat Prot Dosimetry. 168(3): 330-336.

IAEA. 2001. Cytogenetic Analysis for Radiation Dose Assessment, Technical Reports Series No. 405. Vienna, International Atomic Energy Agency.

IAEA. 2011. Cytogenetic Dosimetry: Applications in Preparedness for and Response to Radiation Emergencies, EPR Biodosimetry. Vienna, International Atomic Energy Agency.

ISO. 21243. 2008. Radiation protection - Performance criteria for laboratories performing cytogenetic triage for assessment of mass casualties in radiological or nuclear emergencies -- General principles and application to dicentric assay. Geneva, ISO

ISO 19238. 2014. Radiological protection - Performance criteria for service laboratories performing biological dosimetry by cytogenetics. Geneva, ISO.

ISO 20046. 2019. Radiological protection - Performance criteria for laboratories using Fluorescence In Situ Hybridization (FISH) translocation assay for assessment of exposure to ionizing radiation. Geneva, ISO.

Jackson CH. 2011. Multi-State Models for Panel Data: The msm Package for R. J Stat Softw. 38(8): 1–29.

Kulka U, Abend M, Ainsbury E, Badie C, Barquinero JF, Barrios L, Beinke C, Bortolin E, Cucu, A, de Amicis A, et al. 2017. RENEB - Running the European Network of biological dosimetry and physical retrospective dosimetry. Int J Radiat Biol. 93(1): 2–14.

Lamport L. 1994. LaTeX: A document preparation system (2nd ed.). Addison-Wesley.

https://www.latex-project.org/

Lea DE, Catcheside DG. 1942. The mechanism of the induction by radiation of chromosome aberrations in Tradescantia. J Genet. 44(2–3): 216–245.

Lloyd DC, Edwards AA. 1983. Chromosome aberrations in human lymphocytes: Effect of radiation quality, dose, and dose rate. In: Ishihara T, Sasaki M, editors. Radiation-Induced Chromosome Damage in Man. New York (NY): Alan R. Lis, Inc. p. 23–49.

Lloyd DC, Edwards AA, Prosser JS, Corp MJ. 1984. The dose-effect relationship obtained at constant irradiation times for the induction of chromosome aberrations in human lymphocytes by cobalt-60 gamma rays. Radiat Environ Biophys. 23(3): 179–189.

Lloyd DC, Purrott RJ, Reeder EJ. 1980. The incidence of unstable chromosome aberrations in peripheral blood lymphocytes from unirradiated and occupationally exposed people. Mutat. Res. - Fundam. Mol. Mech. Mutagen. 72(3): 523–532.

Lucas JN, Awa A, Straume T, Poggensee M, Kodama Y, Nakano M, Ohtaki K, Weier H-U, Pinkel D, Gray J, et al. 1992. Rapid Translocation Frequency Analysis in Humans Decades after Exposure to Ionizing Radiation. Int J Radiat Biol. 62(1): 53–63.

Lucas JN, Deng W. 2000. Views on Issues in Radiation Biodosimetry Based on Chromosome Translocations Measured by FISH. Radiat Prot Dosimetry. 88(1): 77–86.

Lyttle I. 2021. bsplus: Adds Functionality to the R Markdown + Shiny Bootstrap Framework. https://github.com/ijlyttle/bsplus

MacFarlane J. 2022. Pandoc: A universal document converter (2.19.2). https://pandoc.org/

Merkle W. 1983. Statistical methods in regression and calibration analysis of chromosome aberration data. Radiat Environ Biophys. 21(3): 217–233.

Moriña D, Higueras M, Puig P, Ainsbury EA, Rothkamm K. 2015. radir package: an R implementation for cytogenetic biodosimetry dose estimation. J Radiol Prot. 35(3): 557–569.

Morton NE. 1991. Parameters of the human genome. Proc Natl Acad Sci U S A. 88(17): 7474–7476.

Oliveira M, Einbeck J, Higueras M, Ainsbury EA, Puig P, Rothkamm K. 2016. Zero-inflated regression models for radiation-induced chromosome aberration data: A comparative study. Biom J. 58(2): 259–279.

Owen J. 2021. rhandsontable: Interface to the Handsontable.js Library.

http://jrowen.github.io/rhandsontable/

Papworth DG. 1975. Curve fitting by maximum likelihood. Radiat Bot. 15, 127–140.

Perrier V, Meyer F, Granjon D. 2022. shinyWidgets: Custom Inputs Widgets for Shiny. https://github.com/dreamRs/shinyWidgets

Port M. 2021. RENEB inter-laboratory study on biological and physical dosimetry employing eight assays. Radiat Res. Society's 67th Annual Meeting.

Puig P, Barquinero JF. 2011. An application of compound Poisson modelling to biological dosimetry. Proc Math Phys Eng Sci. 467: 897–910.

Pujol M, Barquinero JF, Puig P, Puig R, Caballín MR, Barrios L. 2014. A New Model of Biodosimetry to Integrate Low and High Doses. PLoS One. 2;9(12):e114137

Pujol M, Barrios L, Puig P, Caballín MR, Barquinero JF. 2016. A New Model for Biological Dose Assessment in Cases of Heterogeneous Exposures to Ionizing Radiation. Radiat Res., 185(2): 151–162.

Pujol-Canadell M, Puig R, Armengol G, Barrios L, Barquinero JF. 2020. Chromosomal aberration dynamics through the cell cycle. DNA Repair (Amst). 89:102838.

R Core Team. (2022). R: A Language and Environment for Statistical Computing. https://www.r-project.org/

Rao CR, Chakravarti IM. 1956. Some Small Sample Tests of Significance for a Poisson Distribution. Biometrics, 12(3), 264-282.

Sasaki MS. 2003. Chromosomal biodosimetry by unfolding a mixed Poisson distribution: a generalized model. Int J Radiat Biol. 79(2): 83–97.

Sasaki MS, Miyata H. 1968. Biological Dosimetry in Atomic Bomb Survivors. Nature, 220 (5173): 1189–1193.

Savage JRK. 1970. Sites of radiation induced chromosome exchanges. Curr. Top. Radiat. Res. 6: 129–194.

Savage JRK, Papworth DG, Bauchinger M, Natarajan AT, Pantelias GE, Griffin CS, Figgitt M, Knehr S, Braselmann H, Darroudi F, et al. 2000. Constructing a 2B Calibration Curve for Retrospective Dose Reconstruction. Radiat Prot Dosimetry. 88(1): 69–76.

Sigurdson AJ, Ha M, Hauptmann M., Bhatti P, Sram RJ, Beskid O, Tawn EJ, Whitehouse CA,

Lindholm C, Nakano M, Kodama Y, et al. 2008. International study of factors affecting human chromosome translocations. Mutat Res Genet Toxicol Environ Mutagen. 652(2): 112–121.

Szłuińska M, Edwards A, Lloyd D. 2007. Presenting statistical uncertainty on cytogenetic dose estimates. Radiat Prot Dosimetry. 123(4): 443–449.

Trompier F, Baumann M, Barrios L, Gregoire E, Abend M, Ainsbury E, Barnard S, Barquinero JF, Bautista JA, Brzozowska B, et al. 2017. Investigation of the influence of calibration practices on cytogenetic laboratory performance for dose estimation. Int J Radiat Biol. 93(1): 118–126.

Venables WN, Ripley BD. 2002. Modern Applied Statistics with S (4th ed.). Springer. https://www.stats.ox.ac.uk/pub/MASS4/

Virsik RP, Harder D. 1981. Statistical Interpretation of the Overdispersed Distribution of Radiation-Induced Dicentric Chromosome Aberrations at High LET. Radiat Res. 85(1): 13–23.

Vral A, Fenech M, Thierens H. 2011. The micronucleus assay as a biological dosimeter of in vivo ionising radiation exposure. Mutagenesis, 26(1): 11–17.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer International Publishing.

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al. 2019. Welcome to the Tidyverse. J Open Source Softw. 4(43): 1686.

Wickham H, François R, Henry L, Müller K. 2022. dplyr: A Grammar of Data Manipulation. https://dplyr.tidyverse.org

Wickham H, Girlich M. 2022. tidyr: Tidy Messy Data. https://tidyr.tidyverse.org

Table 1. List of historical analysis software developed specifically for radiation biodosimetry.

Name	Language (s)	Framework(s)	Author (s)
MLPOL	Fortran 77	-	(Edwards 1994)
MLREG	-	-	(BfS 1996)
DOSGEN	Pascal	Turbo Pascal	(Garcia Lima and Tomas Zerquera 1996))
CABAS	Object Pascal	Delphi	(Deperas et al. 2007)
Dose Estimate	Object Pascal	Delphi 6	(Ainsbury and Lloyd 2010)

Table 2. Dicentric distribution within cells, dispersion index, and *u*-value of the dose-effect curve constructed from cytogenetic results obtained from blood samples irradiated with γ-rays (Barquinero et al. 1995).

D (Gy)	*N*	*X*	C_0	C_1	C_2	C_3	C_4	C_5	$\bar{y}$	$\hat{\sigma}^2$	$\hat{\sigma}^2/\bar{y}$	*u*
0.00	5000	8	4992	8	0	0	0	0	0.002	0.002	0.999	-0.075
0.10	5002	14	4988	14	0	0	0	0	0.003	0.003	0.997	-0.135
0.25	2008	22	1987	20	1	0	0	0	0.011	0.012	1.080	2.610
0.50	2002	55	1947	55	0	0	0	0	0.027	0.027	0.973	-0.861
0.75	1832	100	1736	92	4	0	0	0	0.055	0.056	1.026	0.790
1.00	1168	109	1064	99	5	0	0	0	0.093	0.093	0.999	-0.018
1.50	562	100	474	76	12	0	0	0	0.178	0.189	1.064	1.077
2.00	333	103	251	63	17	2	0	0	0.309	0.353	1.141	1.822
3.00	193	108	104	72	15	2	0	0	0.560	0.466	0.834	-1.638
4.00	103	103	35	41	21	4	2	0	1.000	0.882	0.882	-0.844
5.00	59	107	11	19	11	9	6	3	1.814	2.085	1.150	0.811

For each dose analyzed, the total number of cells scored (*N*), the total number of dicentrics observed (*X*), the cell distribution of dicentrics ($C_{0,...}$ C_5), and the dispersion index ($\hat{\sigma}^2/\bar{y}$) and *u*-test statistic (*u*) are presented (a sample is considered overdispersed if u > 1.96).

Table 3. Fitted values of the coefficients of the LQ dose-effect calibration curve constructed from the dicentric distribution in Table 2.

Coefficient	Estimate	SE	*t*-statistic	*p*-value
C	0.00128	0.00047	2.716	0.007
α	0.02104	0.00516	4.079	<0.001
β	0.06303	0.00401	15.730	<0.001

Table 4. List of available functions in the biodosetools package. The comprehensive functions reference can be found on the project's website at https://biodosetools-team.github.io/biodosetools/reference/.

Usage	Functions
Fitting	fit(), plot_fit_dose_curve()
Estimation	estimate_whole_body_merkle(), estimate_whole_body_ delta(), estimate_partial_body_dolphin(), estimate_ hetero_mixed_poisson(), plot_estimated_dose_curve()
Auxiliary	calculate_aberr_table(), protracted_g_function(), calculate_genome_fraction(), calculate_trans_rate_ manual(), calculate_trans_rate_sigurdson()
shiny app	run_app()

Table 5. Example of a calibration data.

D (Gy)	*N*	*X*
0.00	100	0
0.25	200	1
0.50	161	6
1.00	88	10
1.50	131	28
2.00	74	23
2.50	189	70
3.00	141	61

D, dose in Gy. Number of cells analyzed, and X number of dicentrics observed.

Table 6. Example of dicentric distribution within cells.

N	X	C_0	C_1	C_2	$\bar{y}$	SE	$\hat{\sigma}^2/\bar{y}$	u
148	13	136	11	1	0.088	0.025	1.073	0.654

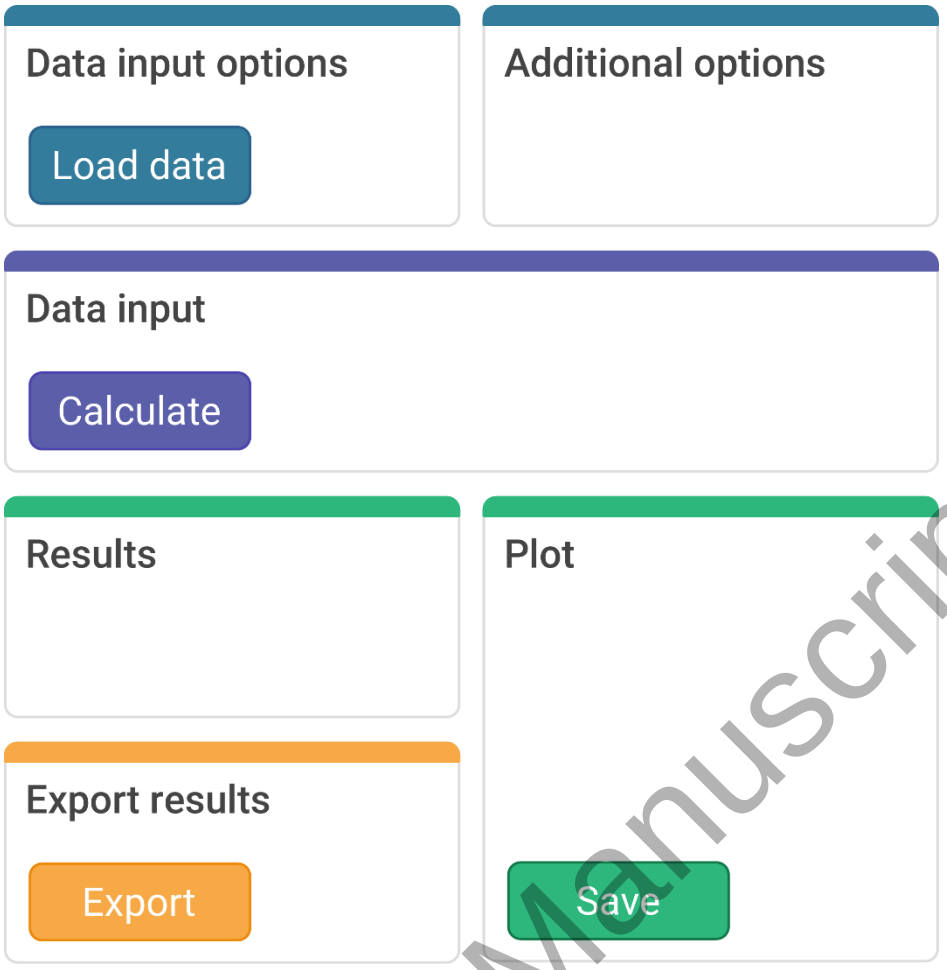


Figure 1. Mock-up illustrating Biodose Tools' color coding to create an intuitive visual organization of the user interface.

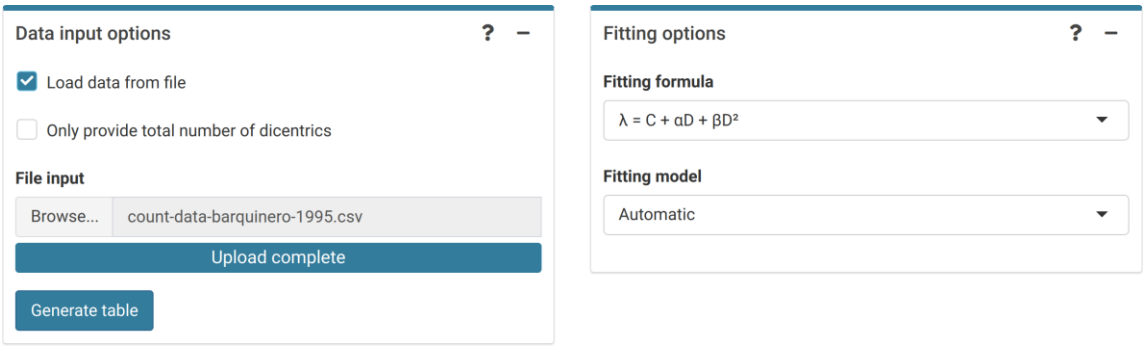


Figure 2. "Data input options" and "Fitting options" boxes in the dose-effect fitting module. For dicentrics, the "Automatic" fitting model will select a quasi-Poisson model if there is overdispersion on the fitting, otherwise it will select a Poisson model assuming equidispersion.

Data input

	D (Gy)	N	X	C0	C1	C2	C3	C4	C5	mean	var	DI	u
1	0.000	5000	8	4992	8	0	0	0	0	0.00	0.00	1.00	-0.07
2	0.100	5002	14	4988	14	0	0	0	0	0.00	0.00	1.00	-0.13
3	0.250	2008	22	1987	20	1	0	0	0	0.01	0.01	1.08	2.61
4	0.500	2002	55	1947	55	0	0	0	0	0.03	0.03	0.97	-0.86
5	0.750	1832	100	1736	92	4	0	0	0	0.05	0.06	1.03	0.79
6	1.000	1168	109	1064	99	5	0	0	0	0.09	0.09	1.00	-0.02
7	1.500	562	100	474	76	12	0	0	0	0.18	0.19	1.06	1.08
8	2.000	333	103	251	63	17	2	0	0	0.31	0.35	1.14	1.82
9	3.000	193	108	104	72	15	2	0	0	0.56	0.47	0.83	-1.64
10	4.000	103	103	35	41	21	4	2	0	1.00	0.88	0.88	-0.84
11	5.000	59	107	11	19	11	9	6	3	1.81	2.09	1.15	0.81

Calculate parameters | Save count data | .csv | Calculate fitting

Figure 3. "Data input" box in the dose-effect fitting module.

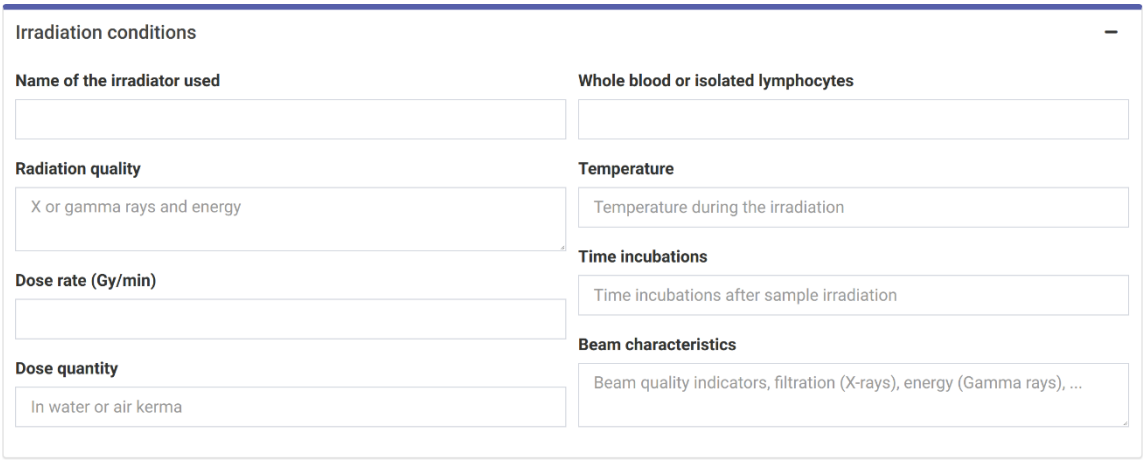


Figure 4. "Irradiation conditions" box in the dose-effect fitting module.

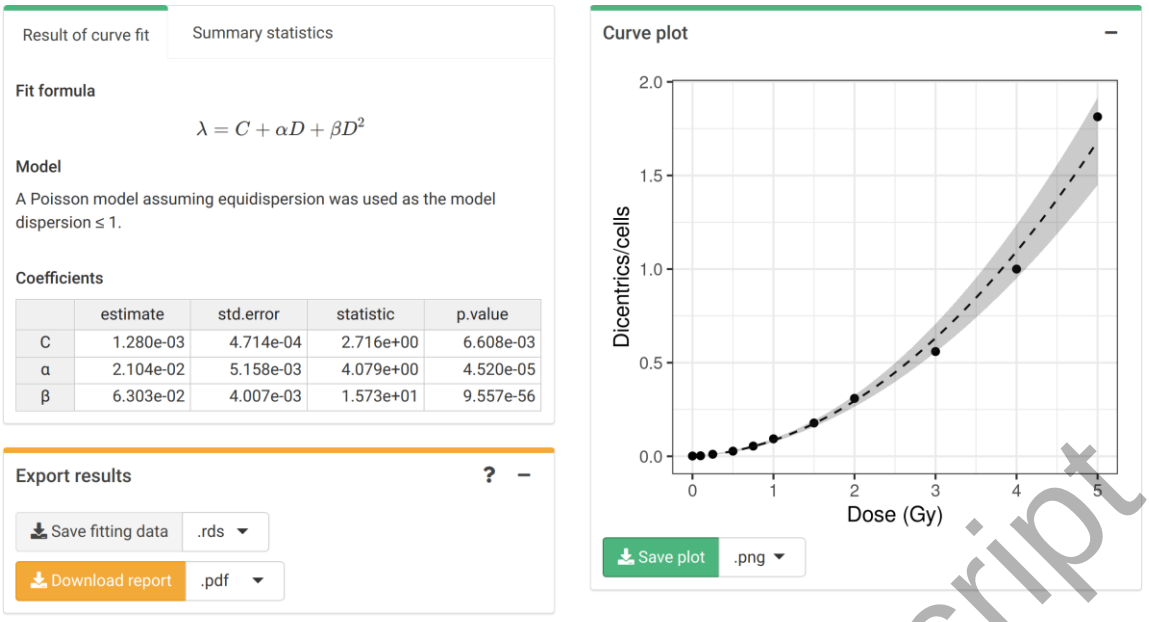


Figure 5. "Results" tabbed box, "Curve plot" and "Export results" boxes in the dose-effect fitting module.

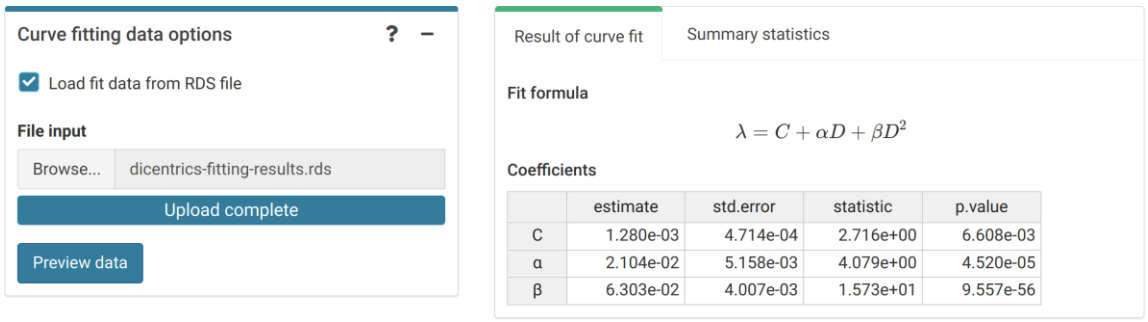


Figure 6. "Curve fitting data options" box and "Results" tabbed box in the dose estimation module when loading curve from an .rds file.

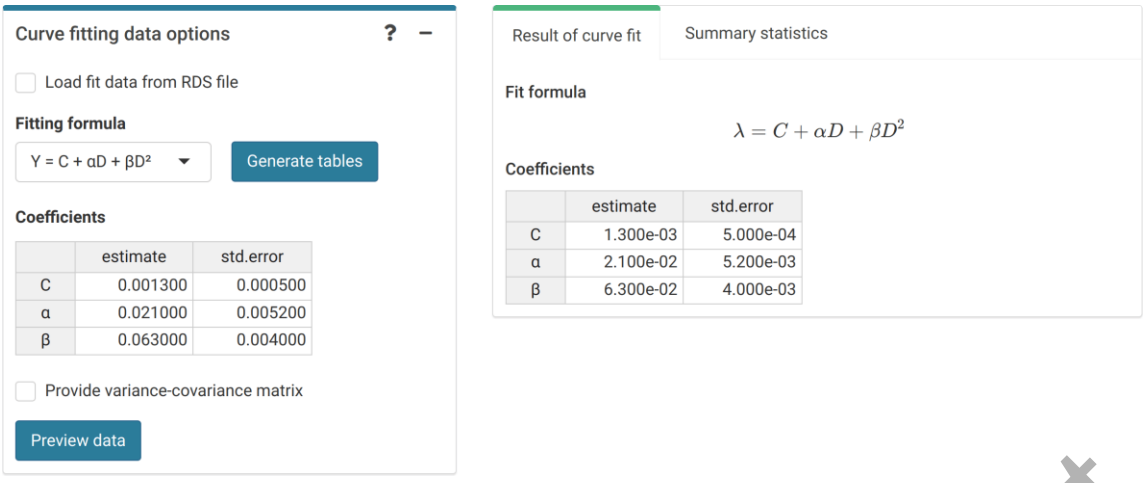


Figure 7. "Curve fitting data options" box and "Results" tabbed box in the dose estimation module when inputting curve coefficients manually. Note that if no variance-covariance matrix is provided, only the variances calculated from the coefficients' standard errors will be used in equations (S1.6), (S1.9), and (S1.19).

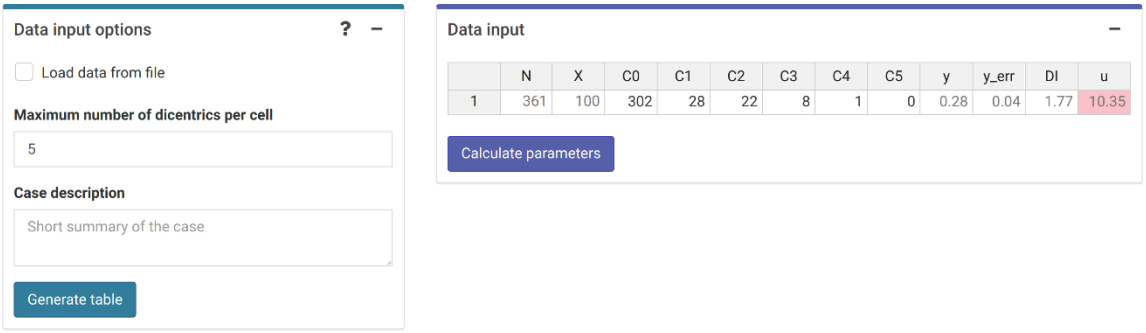


Figure 8. "Data input options" and "Data input" boxes in the dose estimation module.

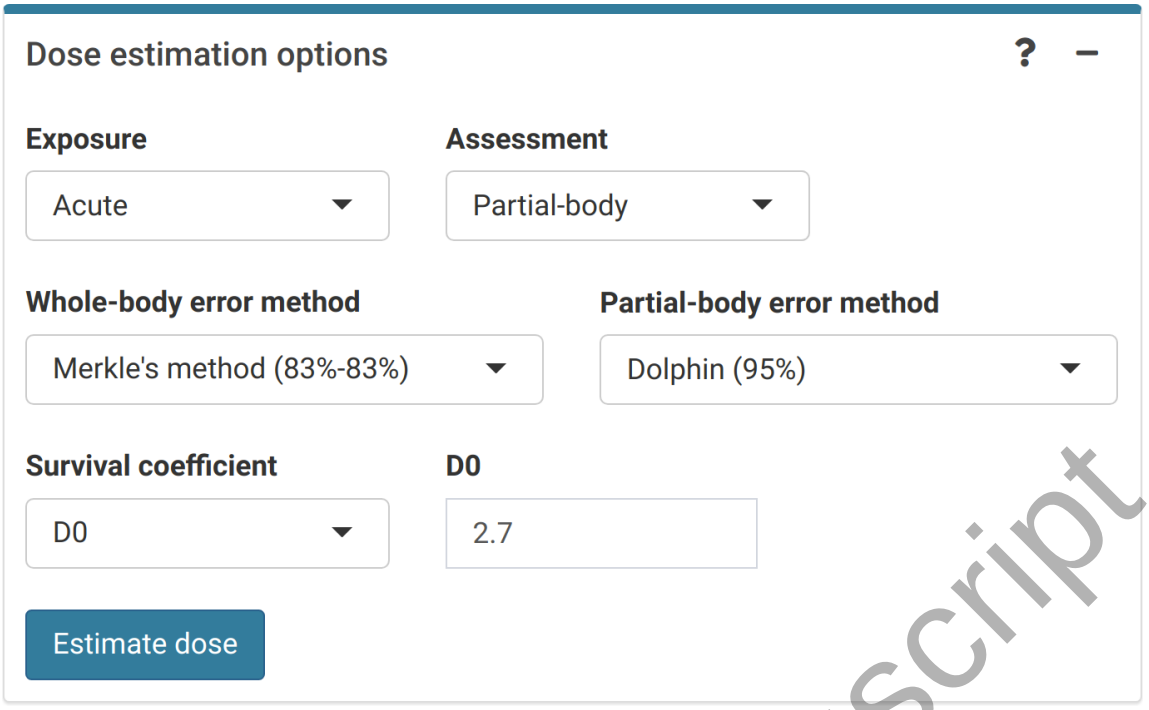


Figure 9. "Dose estimation options" box in the dose estimation module.

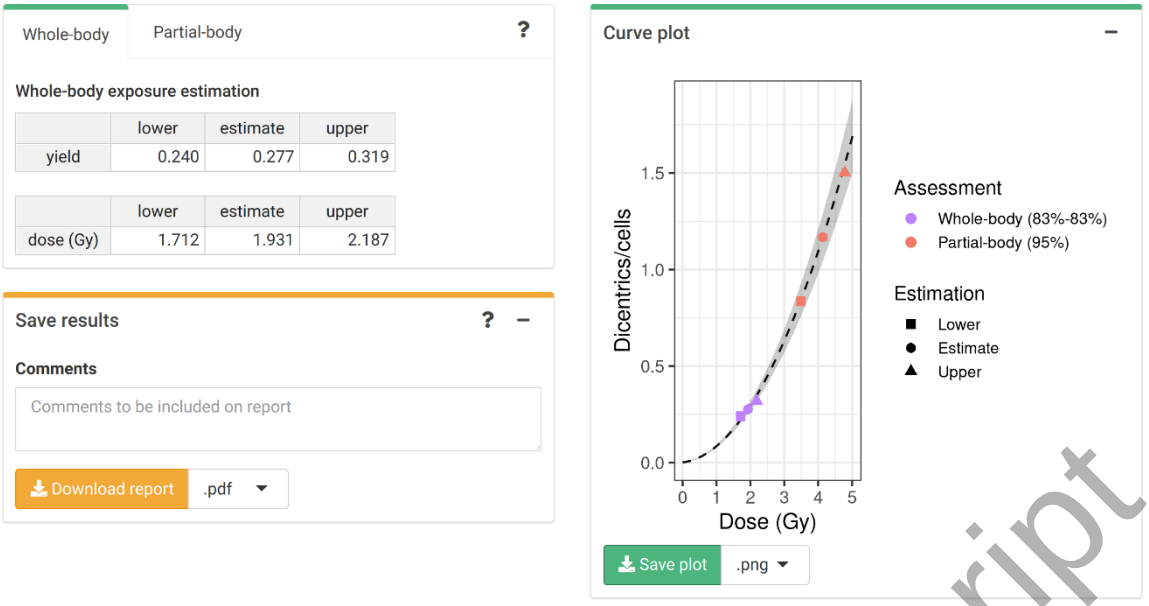


Figure 10. "Results" tabbed box, "Curve plot" and "Save results" boxes in the dose estimation module.

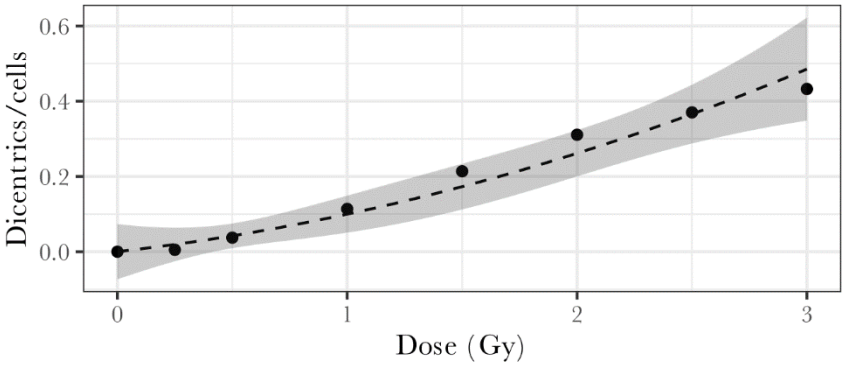


Figure 11. Plot of dose-effect curve constructed from the dicentric distribution in Table 5.

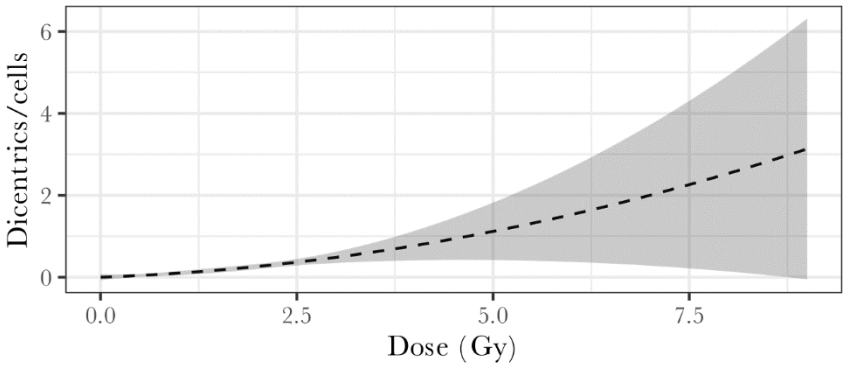


Figure 12. Plot of dose-effect curve constructed from the dicentric distribution in Table 5 over an extended *D* range.