Public Health England

Protecting and improving the nation's health

Bayesian solutions to biodosimetry count data problems, and software solutions

M Higueras, P Puig, EA Ainsbury, D Moriña, VA Vinnikov and K Rothkamm

Review

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. Dosim.* **162**(3), 185-196.

The Bayesian approach

A Bayesian approach is highly applicable to ionising radiation dosimetry data. It allows the cytogenetic experts to consider prior knowledge surrounding an overexposure scenario.

This approach implies an accurate measure of the uncertainty of dose estimates.

The calibrative density is the solution of the Bayesian inverse regression problem,

$$\mathrm{P}(x|y) \propto \mathrm{P}(x) \int \mathrm{L}(y|\Theta)\mathrm{P}(\Theta|x)\mathrm{d}\Theta.$$

Groer & Pereira (1987)

Poisson responses and just linear dose-response (without intercept).

$$Y_i \sim \text{Pois}(\alpha d_i),\ \alpha \sim \text{Gamma}(a, b) \Longrightarrow \alpha|Y \sim \text{Gamma}(a + S, b + N),$$

S is the total number of aberrations and

$$N = \sum n_i d_i.$$

The absorbed dose

$$D \sim \text{Gamma}(A, B),$$

and the calibrative dose density results

$$\text{p}(D|Y) \propto \frac{D^{A+s-1}}{\text{e}^{BD}(nD + a + N)^{S+s+b}}.$$

Brame & Groer (2002)

NB responses with just linear dose-response (without intercept)

$$Y_i \sim \mathrm{NB}(\alpha d_i, \Psi),\ \Psi \sim \mathrm{Gamma},\ \mathrm{P}(\alpha, \Psi) = \mathrm{P}(\alpha)\mathrm{P}(\Psi)$$

and the priors for ***α*** are Uniform or Normal.

The NB model is compared to the Poisson one using the Bayes Factor

$$BF = \frac{\int \mathrm{L}_{\mathrm{NB}}(Y|\alpha, \Psi)\mathrm{P}(\alpha)\mathrm{P}(\Psi)\mathrm{d}\alpha\mathrm{d}\Psi}{\int \mathrm{L}_{\mathrm{Pois}}(Y|\alpha)\mathrm{P}(\alpha)\mathrm{d}\alpha}.$$

The calibrative dose density results

$$\mathrm{P}(D|Y) \propto \mathrm{P}(D) \int \frac{\Gamma(\Psi^{-1}+s)}{\Gamma(s+1)\Gamma(\Psi^{-1})} \left(\frac{\Psi n \alpha D}{1+\Psi n \alpha D}\right)^s \left(\frac{1}{1+\Psi n \alpha D}\right)^{1/\Psi} \mathrm{P}(\alpha, \Psi|Y)\mathrm{d}\alpha\mathrm{d}\Psi.$$

The integrals in this methodology are done using numerical integration.

Other Bayesian works in biodosimetry

- **Madruga *et al.* 1994, 1996:** log-normal model.
- **Kottas *et al.* 2002:** non-parametric model.
- **Serna *et al.* 2008:** Jeffrey's prior for analyzing the background distribution.

Whole body exposure

Higueras M, Puig P, Ainsbury EA, Rothkamm K (2015a). A new Inverse Regression Model Applied to Radiation Biodosimetry. *Proc. R. Soc. A*, DOI: 10.1098/rspa.2014.0588.

Poisson models

The posterior of the population mean is approximated to a normal, by asymptotic normality of the posterior distribution for large samples and the delta-method, i.e.:

$$\mu|D \sim \mathrm{N}\left(\mathrm{f}(D,\hat{\beta}), \nabla \cdot \hat{\Sigma} \cdot \nabla^{\mathrm{T}}\right).$$

The calibrative density results

$$\mathrm{P}(D|Y) \propto \mathrm{P}(D)\mathrm{P}(X_D = s),$$

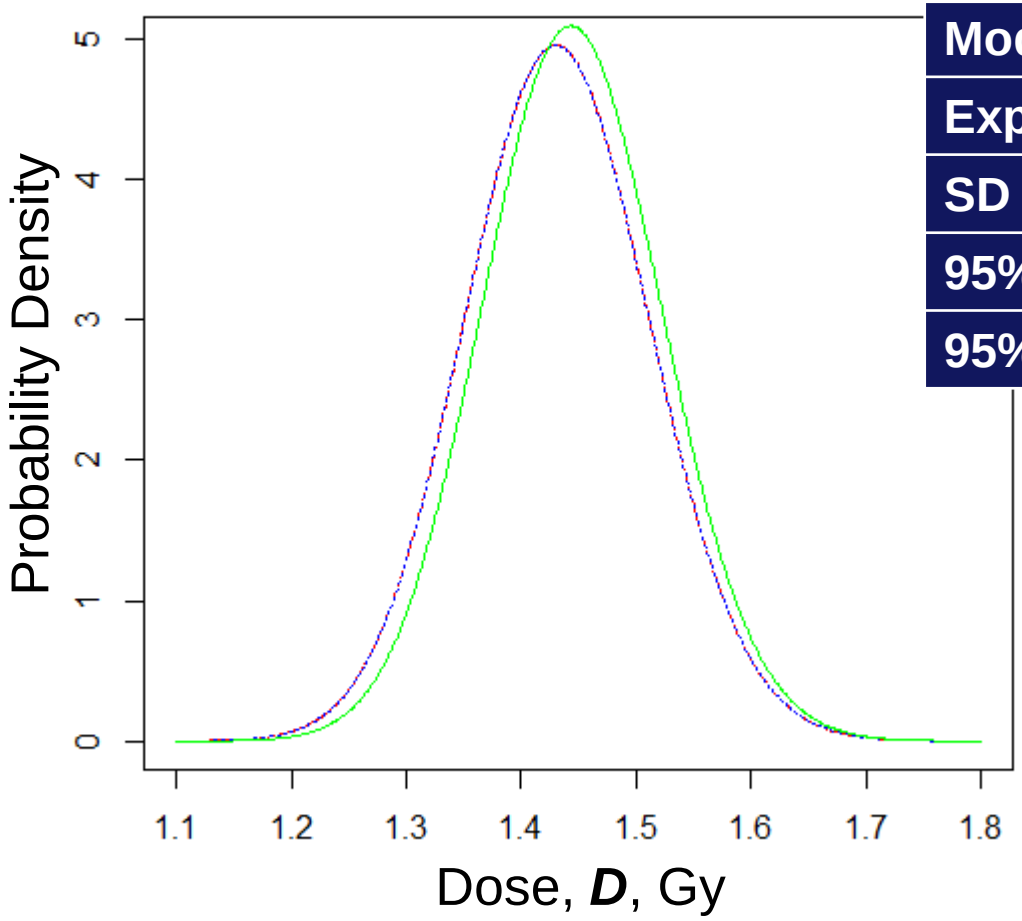
where $\boldsymbol{X_D}$ is Hermite distributed. If $\boldsymbol{\mu|D}$ is approximated by a gamma, $\boldsymbol{X_D}$ is NB distributed.

Example: Romm *et al.* (2013)

- Blood samples from 8 healthy donors were irradiated in vitro with ^{60}Co gamma-rays at a high-dose rate simulating acute whole body exposure.
- Dicentrics assay: Poisson responses and quadratic dose-response (without intercept).
- The 1.5 Gy sample is removed from the calibration dataset to be used as test data.

	Number of dicentrics							
Dose (Gy)	0	1	2	3	4	$\bar{y}$	d	u
0.25	2185	8				0.004	0.997	-0.113
0.75	2550	44	1			0.018	1.026	0.952
1.00	2231	54	2			0.025	1.044	1.503
1.50	*1712*	*96*	*3*			*0.056*	*1.003*	*0.092*
2.50	1196	123	7	1		0.105	1.038	0.985
3.00	1070	320	41	6	1	0.295	1.012	0.334

Example: Romm *et al.* (2013)



Model	(a)	(b)	(c)
Mode	1.430	1.430	1.443
Expected	1.432	1.432	1.445
SD	0.081	0.081	0.078
95% CI LB	1.277	1.277	1.294
95% CI UB	1.594	1.593	1.602

(a): normal mean prior, **U*(0, ∞)*** dose prior.

(b): gamma mean prior, **U*(0, ∞)*** dose prior.

(c): gamma mean prior, **Ga*(21.8, 12.4)*** dose prior.

Compound Poisson models

The joint posterior of the population mean and the dispersion index is defined:

$$(\mu, \delta)|D \sim \mathrm{N}_2\left((\mathrm{f}(D, \hat{\beta}), \hat{\delta}), \nabla \cdot \hat{\Sigma} \cdot \nabla^{\mathrm{T}}\right).$$

The calibrative density can be defined directly and calculated by numerical integration (not computationally intensive, always bivariate: ***D*** and ***δ***).

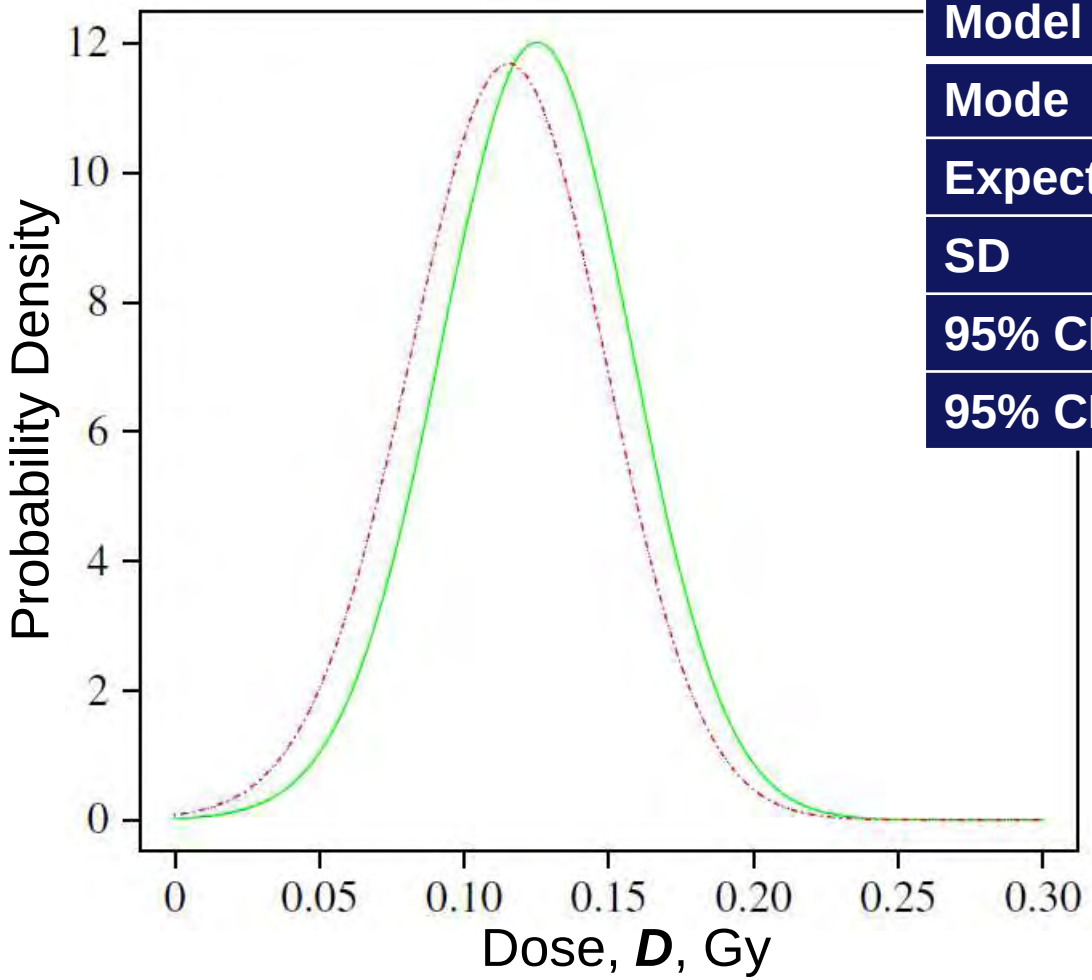
Fixing ***δ*** by its MLE, the model its reduced and the mean prior is applied like in the Poisson models. The resulting calibrative density is in terms of Compound -Hermite (-NB) mass function.

Example: Puig & Valero (2006)

- 11 samples of peripheral blood exposed to different doses of γ-rays (0.93 cGymin^{-1} dose rate). For each sample, approximately 5000 binucleated cells were inspected.
- MN assay: NB responses and quadratic dose-response.
- The 0.1 Gy sample is used as test data.

	no. micronuclei										
dose (Gy)	0	1	2	3	4	5	6	7	$\bar{y}$	d	u
0.00	4887	106	5	2					0.024	1.156	7.839
0.10	*4773*	*206*	*19*	*2*					*0.050*	*1.150*	*7.526*
0.25	4261	324	41	12	2				0.090	1.306	15.306
0.50	4536	364	76	17	7				0.119	1.449	22.484
0.75	4383	512	85	18	2				0.149	1.257	12.876
1.00	4225	636	115	19	5				0.189	1.240	12.009
1.50	4018	805	139	26	9	1	2		0.243	1.270	13.495
2.00	3499	1194	238	45	13	10	1		0.383	1.209	10.471
2.50	3171	1313	393	94	24	3	2		0.501	1.201	10.077
3.00	2582	1575	598	190	44	9	2	6	0.722	1.206	10.307
4.00	1974	1674	869	342	102	26	13	2	1.013	1.172	8.628

Example: Puig & Valero (2006)



Model	(a)	(b)	(c)
Mode	0.115	0.115	0.125
Expected	0.114	0.114	0.124
SD	0.034	0.034	0.033
95% CI LB	0.047	0.047	0.059
95% CI UB	0.182	0.181	0.190

$D \sim U(0, \infty)$

(a): simplified with normal mean prior.

(b): simplified gamma mean prior.

(c): complete.

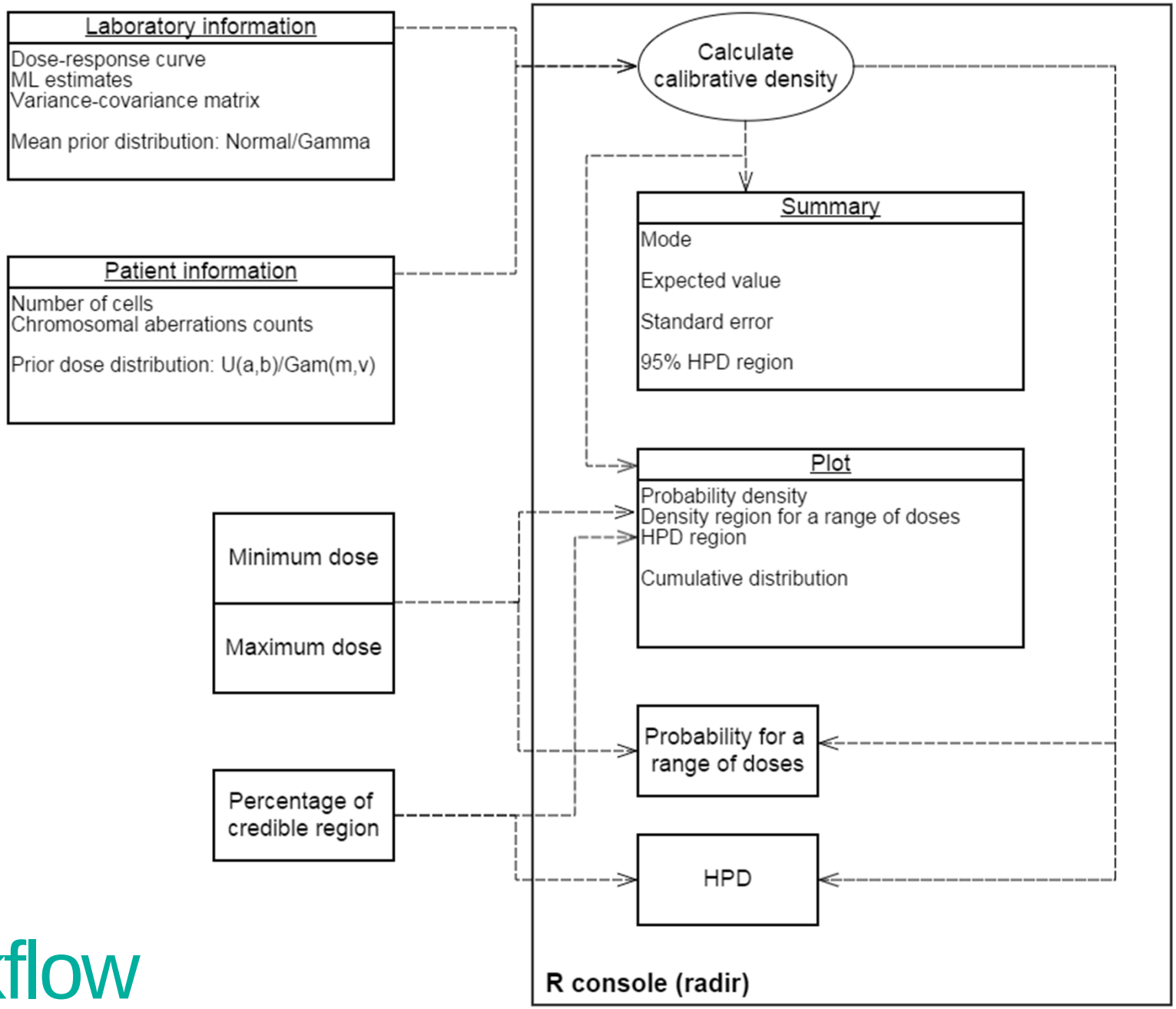
Conclusions of Higueras *et al.* (2015a)

- New whole body cytogenetic dose estimation Bayesian-like models for Poisson and two parameter compound Poisson responses are presented.
- The approach is valid for any given dose-response function one time differentiable in the parameter set domain.
- To use this methodology, only the estimates of the parameters and covariance matrix of the dose-response curve are required.

radir

Moriña D, Higueras M, Puig P, Ainsbury EA, Rothkamm K (2015). `radir` package: An R implementation for cytogenetic biodosimetry dose estimation. *J. Radiol. Prot.*, DOI: 10.1088/0952-4746/35/3/557.

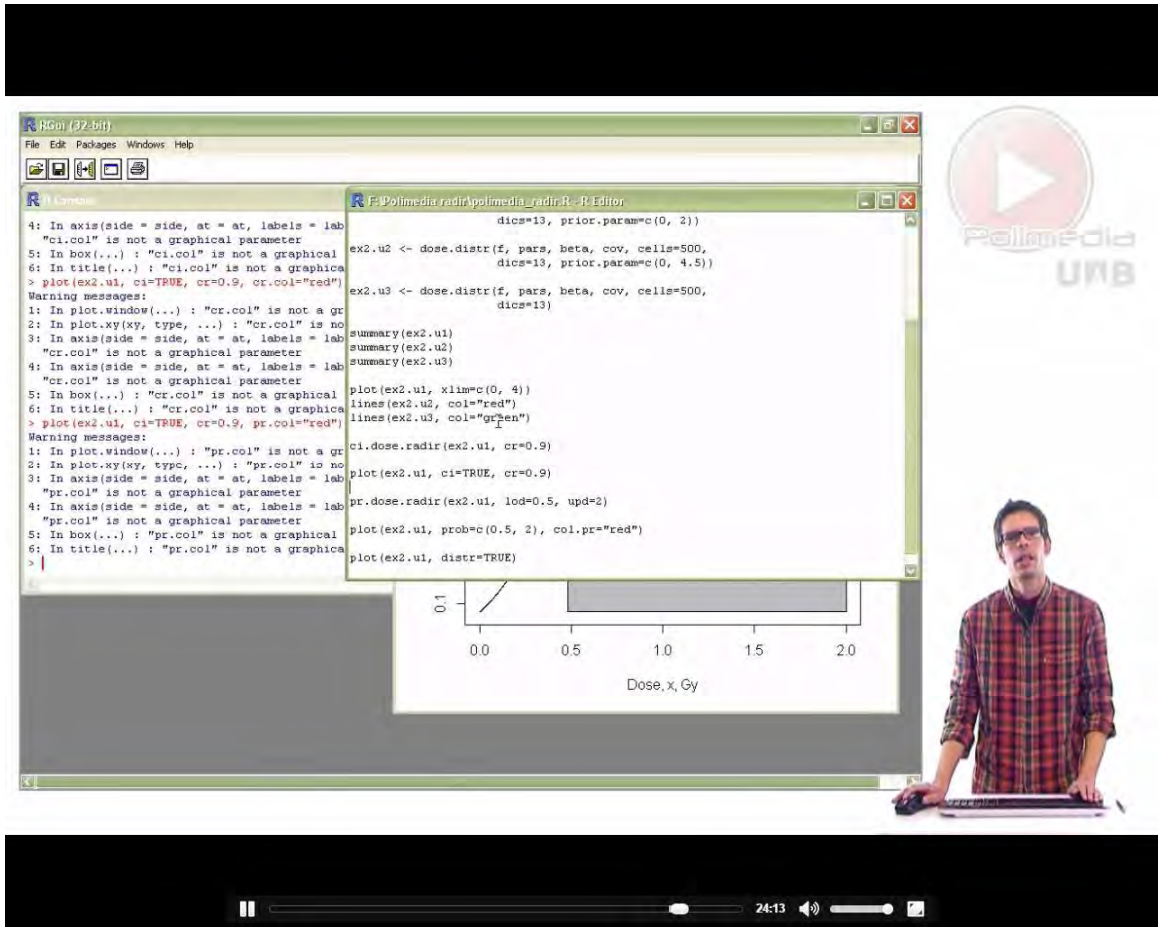
https://cran.r-project.org/web/packages/radir/


Laboratory information
Dose-response curve
ML estimates
Variance-covariance matrix
Mean prior distribution: Normal/Gamma
Patient information
Number of cells
Chromosomal aberrations counts
Prior dose distribution: U(a,b)/Gam(m,v)
Minimum dose
Maximum dose
Percentage of credible region
Calculate calibrative density
Summary
Mode
Expected value
Standard error
95% HPD region
Plot
Probability density
Density region for a range of doses
HPD region
Cumulative distribution
Probability for a range of doses
HPD
R console (radir)

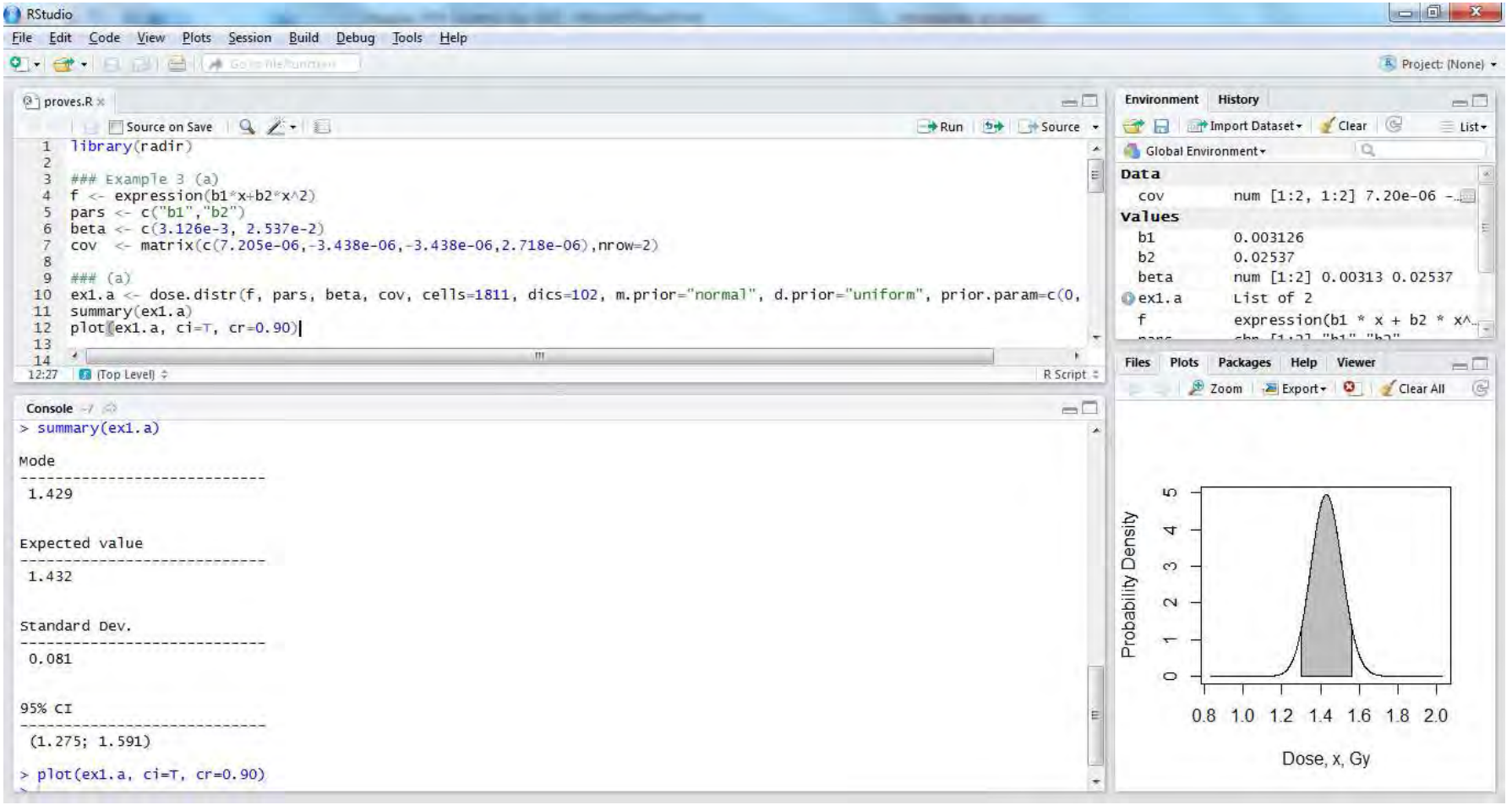
Workflow

Video tutorial

http://polimedia.uab.cat/#v_592



Screenshot


RStudio
File Edit Code View Plots Session Build Debug Tools Help
Project: (None)
proves.R
Source on Save
Run
Source
1 library(radir)
2
3 ### Example 3 (a)
4 f <- expression(b1*x+b2*x^2)
5 pars <- c("b1","b2")
6 beta <- c(3.126e-3, 2.537e-2)
7 cov <- matrix(c(7.205e-06,-3.438e-06,-3.438e-06,2.718e-06),nrow=2)
8
9 ### (a)
10 ex1.a <- dose.distr(f, pars, beta, cov, cells=1811, dics=102, m.prior="normal", d.prior="uniform", prior.param=c(0,
11 summary(ex1.a)
12 plot(ex1.a, ci=T, cr=0.90)
13
14
12:27 (Top Level) R Script
Console
> summary(ex1.a)
Mode
1.429
Expected value
1.432
Standard Dev.
0.081
95% CI
(1.275; 1.591)
> plot(ex1.a, ci=T, cr=0.90)
Environment History
Import Dataset Clear List
Global Environment
Data
cov num [1:2, 1:2] 7.20e-06 -...
Values
b1 0.003126
b2 0.02537
beta num [1:2] 0.00313 0.02537
ex1.a List of 2
f expression(b1 * x + b2 * x^...
Files Plots Packages Help Viewer
Zoom Export Clear All
Probability Density
0 1 2 3 4 5
0.8 1.0 1.2 1.4 1.6 1.8 2.0
Dose, x, Gy

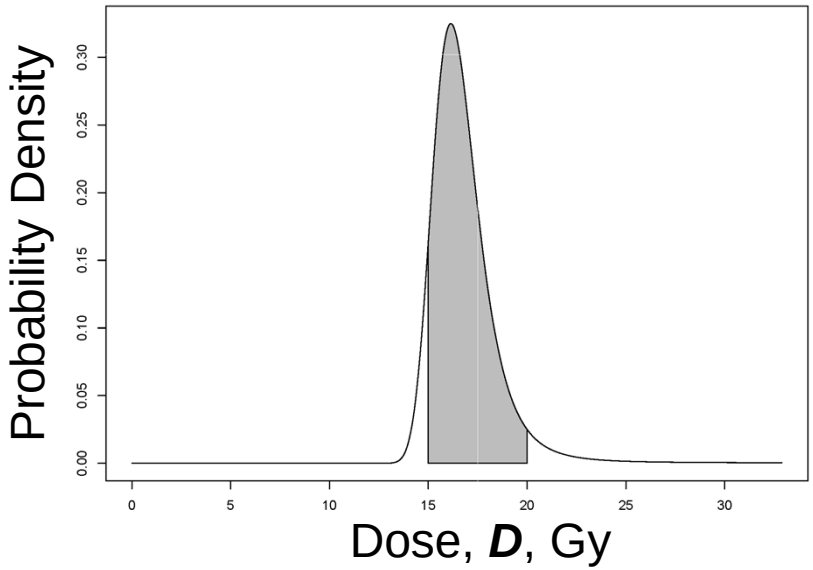
Example: Pujol *et al.* (2014)

The mean of dicentrics per cell as a function of the absorbed dose is

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

We assume Poisson responses. The 17 Gy test sample consists in 914 dicentrics scored in 100 cells. Assuming the gamma mean prior and a ***U(0,∞)*** prior dose.

Statistic	Dose (Gy)
Mode	16.143
Expected	16.814
SD	1.778
95% CI LB	14.148
95% CI UB	19.953



Partial body exposure

Higueras M, Puig P, Ainsbury EA, Vinnikov VA, Rothkamm, K (2015b). A New Bayesian Model Applied to Cytogenetic Partial Body Irradiation Estimation. *Radiat. Prot. Dosim.*, DOI: 10.1098/rspa. 2014.0588.

Bayes factor

The Bayesian alternative for PBI:

- Decision: does the sample come from a partial body exposure? In contrast to the frequentist *u*-test, the Bayarri *et al.* 2008 Bayes factor is proposed (ZIP *vs.* Poisson):

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

  where ***n***, $\boldsymbol{n_0}$ and ***s*** are respectively the sample size and frequency of zeros, and the sum of the total number of chromosomal aberrations.

- Once the ZIP assumption is supported, a new Bayesian model is proposed for the dose and fraction of the body irradiated estimation.

The model

The frequency of aberrations per cell is

$$Z \sim \mathrm{ZIP}\left(\mu, \frac{1-F}{F\mathrm{e}^{-D/d_0} - F + 1}\right),$$

where analogously to Higueras *et al.* (2015a)

$$\mu|D \sim \mathrm{Gamma}\left(\frac{\mathrm{f}(D,\hat{\beta})^2}{\mathrm{v}(D,\hat{\beta})}, \frac{\mathrm{f}(D,\hat{\beta})}{\mathrm{v}(D,\hat{\beta})}\right)$$

an application of Bayes' theorem shows the expression of the likelihood of ***D***, ***F*** and $\boldsymbol{d_0}$ for the given test data

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

where $\boldsymbol{X_j}$ is a random variable negative binomial distribution with mean and variance depending on ***j*** and ***D***.

The model

Considering ***D***, ***F*** and d_0 as independent random variables, and

$$D \sim \text{Gamma}\left(\frac{\hat{D}^2}{\hat{\sigma}^2_{\hat{D}}}, \frac{\hat{D}}{\hat{\sigma}^2_{\hat{D}}}\right);\ F \sim \mathcal{U}(0,1);\ d_0 \sim \mathcal{U}(2.7, 3.5);$$

the joint posterior density,

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

has a non-tractable form. The acceptance-rejection sampling is used to simulate the posterior distribution.

Example

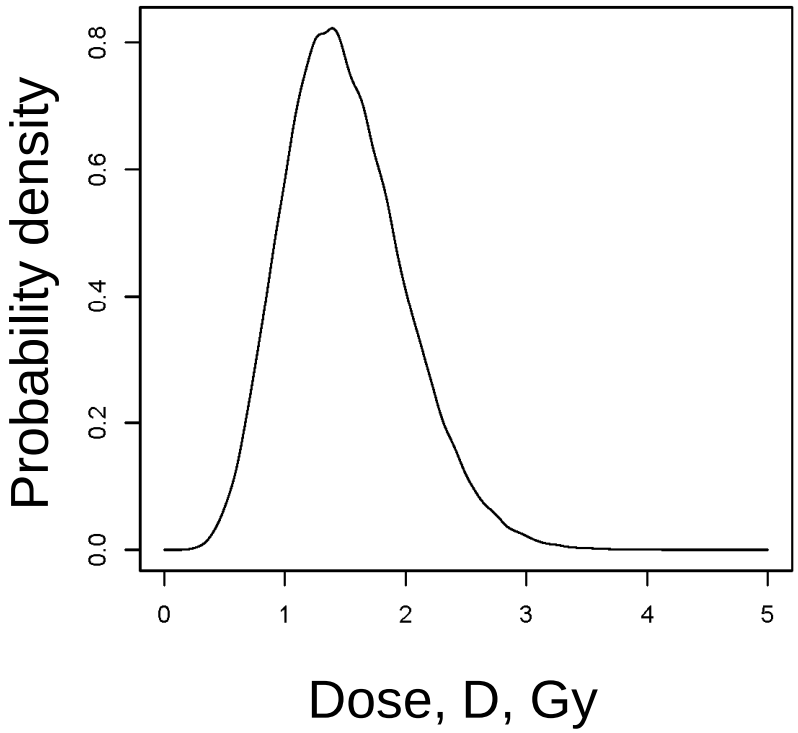
In a recent experiment to simulate PBI, unirradiated and irradiated blood at each dose was mixed. For instance, the exposed blood fraction comprised 10% for 2 Gy, whose distribution of dicentrics plus centric rings is

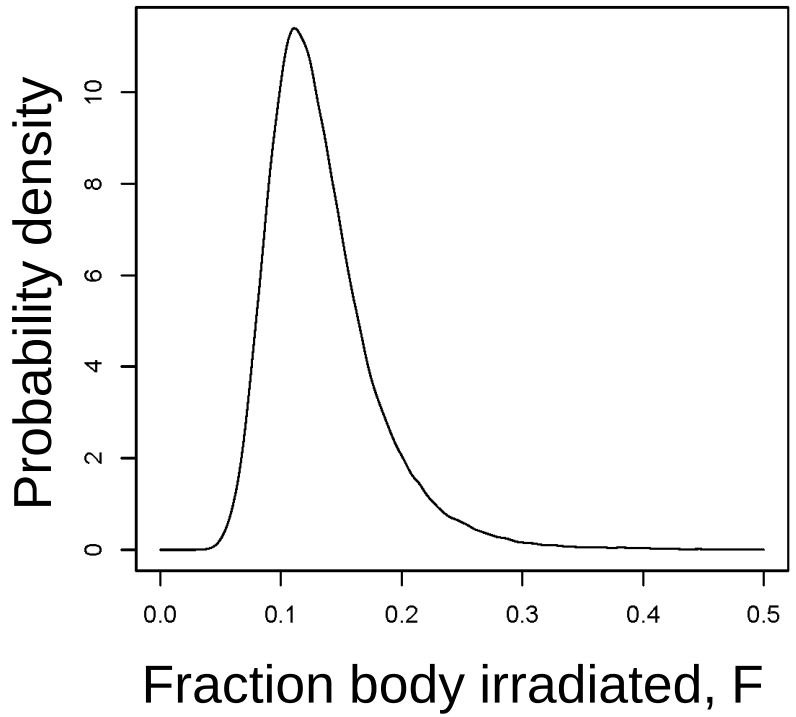
#Dic+CR	0	1	2
#cells	1043	16	3

The Bayes factor value for this sample gives 'strong' evidence in support to the ZIP assumption, because $\boldsymbol{6 < 2 \log BF = 9.41 < 10}$.

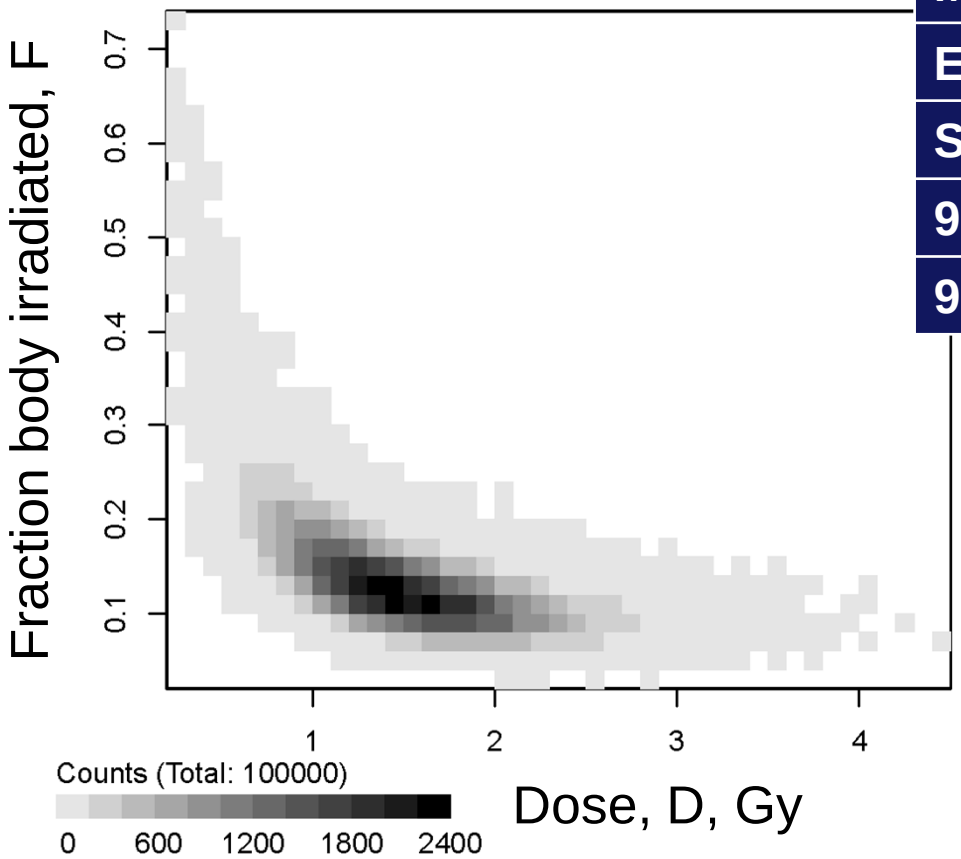
The calibration dose response data is taken from Vinnikov *et al.* 2013.

Marginal posteriors





Joint posterior and statistic summary



Statistic	Dose (Gy)	Fraction
Mode	1.39	0.11
Expected	1.50	0.13
SD	0.50	0.05
95% CI LB	0.66	0.07
95% CI UB	2.58	0.25

Conclusions

- Novel solutions for statistical analysis of cytogenetic biological dosimetry data have been developed.
- New Bayesian models have been created and applied in practical cytogenetic dose estimation.
- Some of these models have been implemented in the `R` statistical software for biodosimetry laboratory researchers.
- These new solutions lead to more accurate quantification of statistical uncertainty associated with cytogenetic dose estimates.

Bibliography

Bayarri MJ, Berger JO and Datta GS (2008). Objective Bayes testing of Poisson versus inflated Poisson models. *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh* 105-121, Institute of Mathematical Statistics, see http://projecteuclid.org/euclid.imsc/1209398464.

Brame RS, Groer PG (2002). Bayesian analysis of overdispersed chromosome aberration data with the negative binomial model. *Radiat. Prot. Dosim.* **102**, 115-119.

Groer PG, Pereira CABD (1987). Calibration of a radiation detector: chromosome dosimetry for neutrons. In: *Probability and Bayesian Statistics*. Ed. R. Viertl, Plenum Publishing Corporation, 225-232.

IAEA (2011). *Cytogenetic Dosimetry: Applications in Preparedness for and Response to Radiation Emergencies*. Vienna: International Atomic Energy Agency, see http://www-pub.iaea.org/MTCD/publications/PDF/ EPRBiodosimetry%202011_web.pdf.

Kottas A, Branco MD, Gelfand AE (2002). A nonparametric Bayesian modelling approach for cytogenetic dosimetry. *Biometrics* **58**, 593-600.

Madruga MR, Pereira CAdB, Rabello-Gay M (1994). Bayesian Dosimetry, Radiation dose versus frequencies of cells with aberrations. *Environmetrics* **5**, 47-56.

Madruga MR, Ochi-Lohmann TH, Okazaki K, Pereira CAdB, Rabello-Gay MN (1996). Bayesian Dosimetry II, Credibility intervals for radiation dose. *Environmetrics* ***7***, 325-331.

Puig P, Valero J (2006). Count data distributions: some characterizations with applications. J. Amer. Statist. Assoc. **101**, 332-340.

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*, DOI: 10.1371/journal.pone.0114137.

Romm H, Ainsbury EA, Barnard S, Barrios L, Barquinero JF, Beinke C, Deperas M, et al (2013). Automatic scoring of dicentric chromosomes as a tool in large scale radiation accidents. *Mutat. Res-Gen. Tox. En.* **756**,174-183.

Serna A, Alcaraz M, Navarro JL, Acevedo C, Vicente V, Canteras M (2008). Biological dosimetry and Bayesian analysis of chromosomal damage in thyroid cancer patients. *Radiat. Prot. Dosim.* **129**, 372-380.

Vinnikov, VA and Maznyk, NA (2013). Cytogenetic dose-response *in vitro* for biological dosimetry after exposure to high doses of gamma-rays. *Radiat. Prot. Dosim.* **154**(2), 186-197.

Acknowledgements

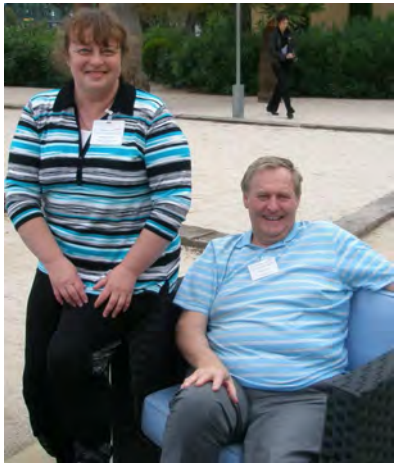
UAB
Universitat Autònoma de Barcelona

Universitätsklinikum Hamburg-Eppendorf

CREAL
centre de recerca en epidemiologia ambiental

Public Health England

Protecting and improving the nation's health



State institution
«Grigoriev Institute for Medical Radiology of National Academy of Medical Science of Ukraine»

...and all the pieces matter.

Thanks for your attention!