Radiation dosimetry through statistical analysis of biomarkers

Jochen Einbeck
jochen.einbeck@durham.ac.uk

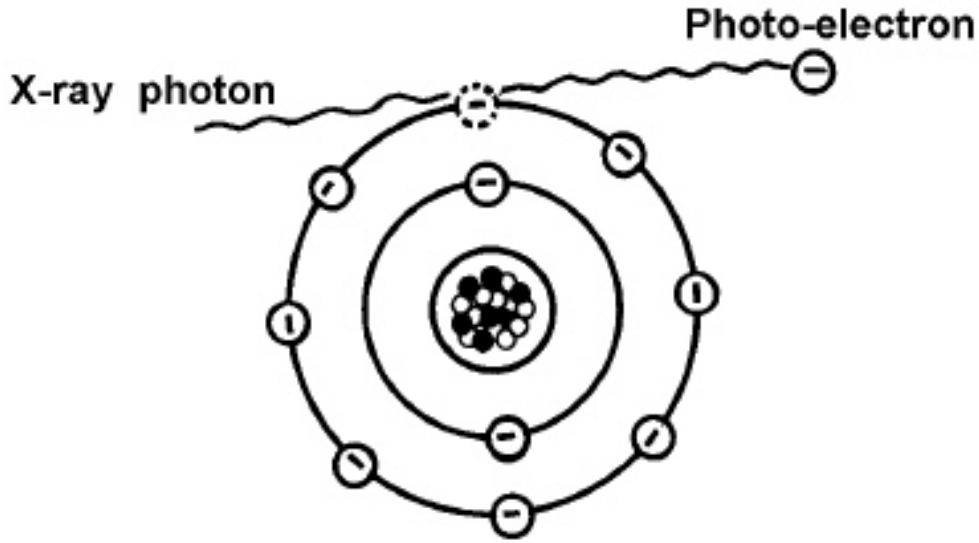
Durham, 9th November 2016

Radiation

- Radiation is energy in the form of waves or particles that travels through space or some material (includes heat, radio waves, light,...)

Radiation

- Radiation is energy in the form of waves or particles that travels through space or some material (includes heat, radio waves, light,...)
- When we talk about radiation, we often mean ionizing radiation (α and β particles, γ–rays, X–rays, neutrons...), which carries enough energy to ionize atoms or molecules. Ionizing radiation can cause serious damage to cells, tissues, and DNA.



Biodosimetry

- Dosimetry is the measurement of the absorbed dose delivered by ionizing radiation. The absorbed dose is measured in Gy (Joules per kg).

Biodosimetry

- Dosimetry is the measurement of the absorbed dose delivered by ionizing radiation. The absorbed dose is measured in Gy (Joules per kg).
- A biomarker (short for: "biological marker") is "any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease" (WHO).
- Biodosimetry is a dosimetry technique which exploits the information provided by radiation–sensitive biomarkers (usually, radiation–induced damage inside the cell nucleus) to infer the radiation dose.

Biodosimetry

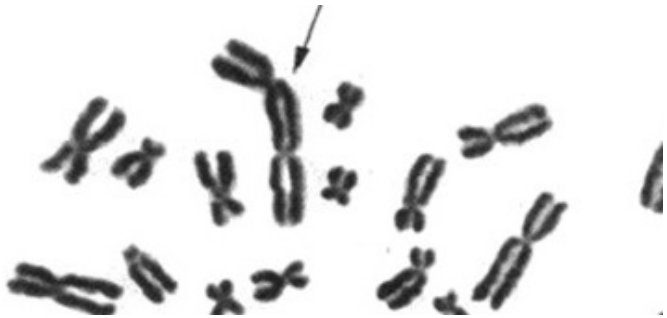
- Dosimetry is the measurement of the absorbed dose delivered by ionizing radiation. The absorbed dose is measured in Gy (Joules per kg).
- A biomarker (short for: "biological marker") is "any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease" (WHO).
- Biodosimetry is a dosimetry technique which exploits the information provided by radiation–sensitive biomarkers (usually, radiation–induced damage inside the cell nucleus) to infer the radiation dose.
- Statistics comes into play when establishing the link between the biomarker and the absorbed dose.

Radiation biodosimetry

- Radiation accident or incident leading to irradiated blood lymphocytes.
- Need rapid and reliable procedures to determine the radiation dose contracted by individuals.
- Members of the public do not usually wear radiation dosimeters...
- Hence, triage and clinical decision making needs to rely on biomarkers to estimate the contracted radiation dose.

Radiation biodosimetry

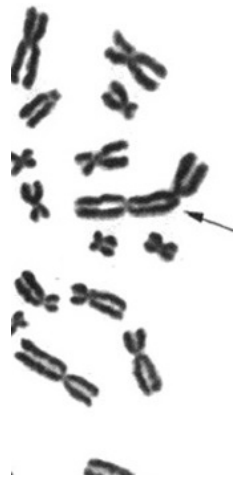
- Radiation accident or incident leading to irradiated blood lymphocytes.
- Need rapid and reliable procedures to determine the radiation dose contracted by individuals.
- Members of the public do not usually wear radiation dosimeters...
- Hence, triage and clinical decision making needs to rely on biomarkers to estimate the contracted radiation dose.
- Most commonly used biomarker: Chromosome aberrations, such as dicentric chromosomes:



Cytogenetic biomarkers

- Example: Frequencies of dicentrics (= aberrant chromosome having two centromeres) in $n = 4400$ lymphocytes after *in vitro* 'whole body' exposure with 200 kV X-rays.

	y_{ij}								
x_i	0	1	2	3	4	5	6	7	n_i
1	1715	268	15	2	0	0	0	0	2000
2	638	298	56	8	0	0	0	0	1000
3	247	225	85	37	6	0	0	0	600
4	99	129	92	52	21	5	2	0	400
5	48	88	97	99	36	25	5	2	400



- x_i: dose (in Gy) used to irradiate blood sample i, $i = 1, \ldots 5$.
- y_{ij}: counts of dicentric aberrations in j-th cell of blood sample i, $j = 1, \ldots n_i$.

Dose–response modelling

- These are count data, so a natural choice for the response distribution is Poisson, that is

$$f(y_{ij}|x_i) = e^{-\lambda_i} \frac{\lambda_i^{y_{ij}}}{y_{ij}!}$$

- There is agreement in the biodosimetry literature that the Poisson means λ_i can be described by a parametric model

$$\lambda_i = E(y_{ij}|x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$$

Dose–response modelling

- These are count data, so a natural choice for the response distribution is Poisson, that is

$$f(y_{ij}|x_i) = e^{-\lambda_i} \frac{\lambda_i^{y_{ij}}}{y_{ij}!}$$

- There is agreement in the biodosimetry literature that the Poisson means λ_i can be described by a parametric model

$$\lambda_i = E(y_{ij}|x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$$

- For parameter estimation, firstly set up the likelihood function:

$$L = \prod_{i,j} f(y_{ij}|x_i) = \prod_{i,j} e^{-\lambda_i} \frac{\lambda_i^{y_{ij}}}{y_{ij}!} \propto \prod_i e^{-n_i \lambda_i} \lambda_i^{\sum_j y_{ij}}$$

Dose–response modelling

- These are count data, so a natural choice for the response distribution is Poisson, that is

$$f(y_{ij}|x_i) = e^{-\lambda_i} \frac{\lambda_i^{y_{ij}}}{y_{ij}!}$$

- There is agreement in the biodosimetry literature that the Poisson means λ_i can be described by a parametric model

$$\lambda_i = E(y_{ij}|x_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$$

- For parameter estimation, firstly set up the likelihood function:

$$L = \prod_{i,j} f(y_{ij}|x_i) = \prod_{i,j} e^{-\lambda_i} \frac{\lambda_i^{y_{ij}}}{y_{ij}!} \propto \prod_i e^{-n_i \lambda_i} \lambda_i^{\sum_j y_{ij}}$$

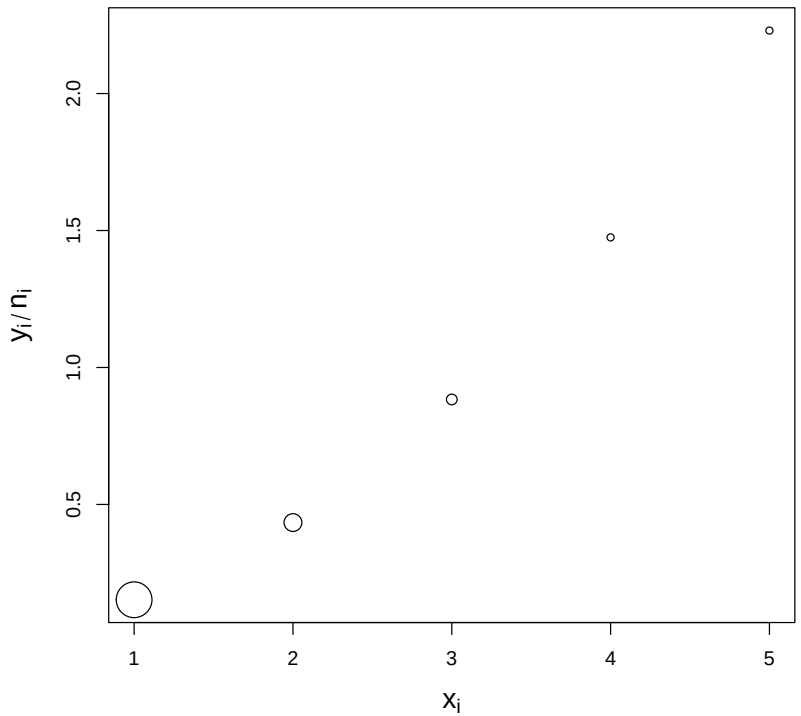
- One can conveniently work at the *aggregated data* level, with data $(x_i, y_i) = (x_i, \sum_{j=1}^{n_i} y_{ij})$.

Aggregated data

- Let $y_i = \sum_j y_{ij}$. Then the aggregated data are

x_i	n_i	y_i
1.0	2000	304
2.0	1000	434
3.0	600	530
4.0	400	590
5.0	400	892

- Graphically, with circle size $\propto n_i$.



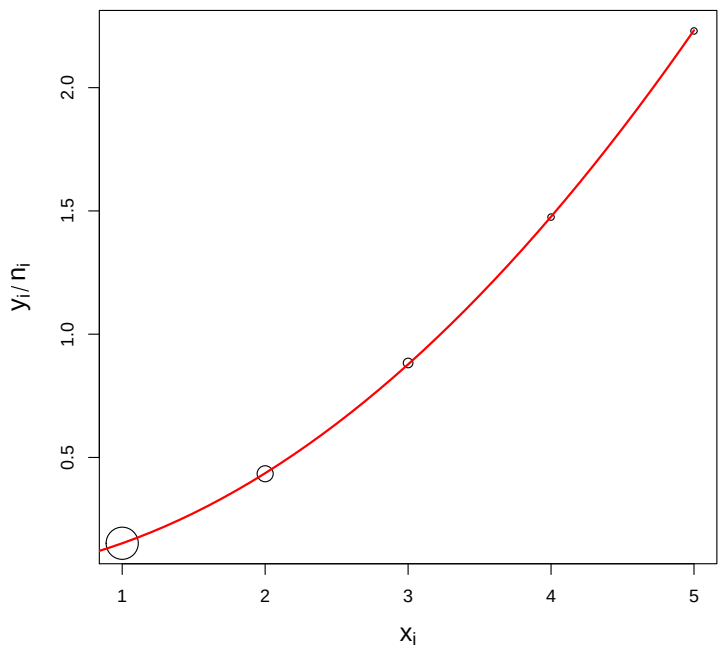
- 'Empirical dose-response curve'

Poisson regression

- Model in terms of aggregated data:

$$\lambda_i \equiv E(y_i|x_i)/n_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$$

  where $y_i \sim \text{Pois}(n_i \lambda_i)$.
- Poisson regression model fitted via ML
  - special case of generalized linear models
- Fitted dose–response curve, $\hat{\lambda}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2$:

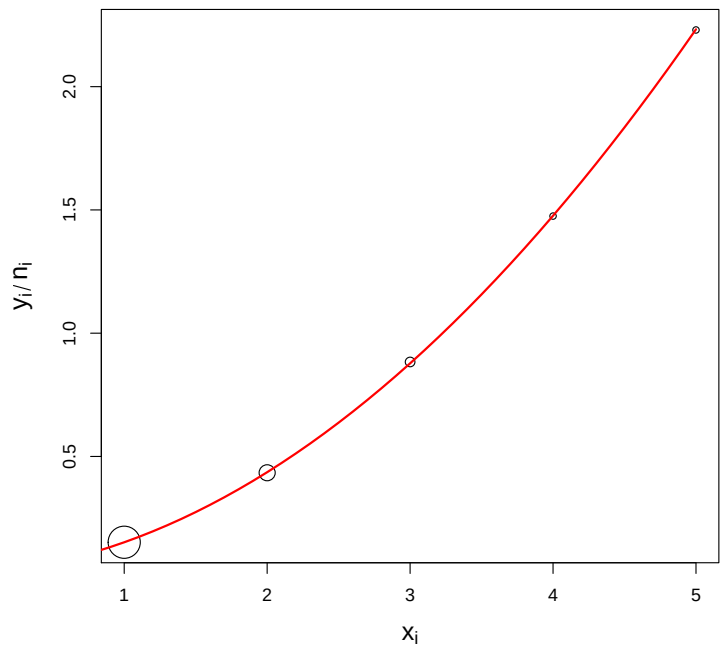


Poisson regression

- Model in terms of aggregated data:

$$\lambda_i \equiv E(y_i|x_i)/n_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$$

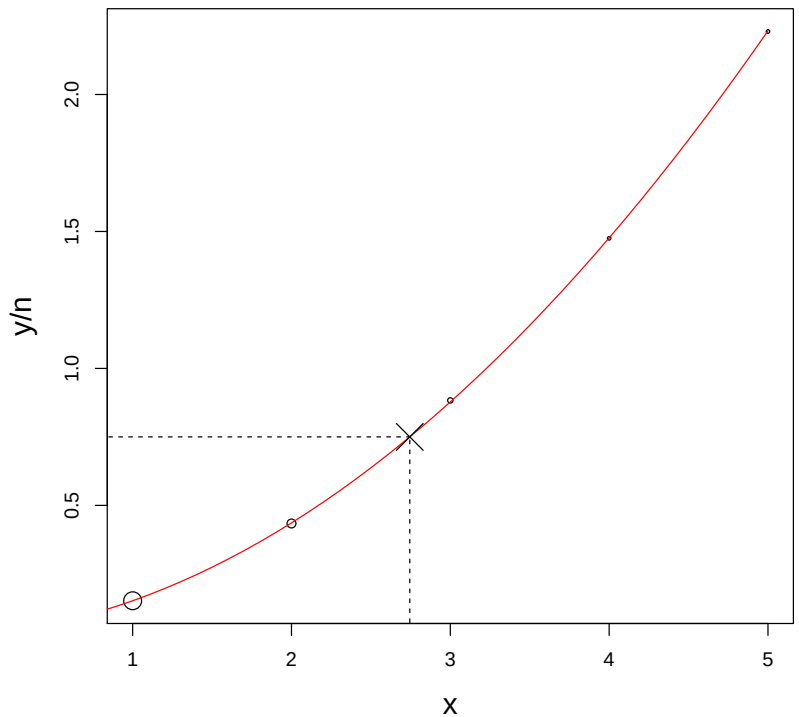
  where $y_i \sim \text{Pois}(n_i \lambda_i)$.
- Poisson regression model fitted via ML
  - special case of generalized linear models
- Fitted dose–response curve, $\hat{\lambda}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2$:



- serves as calibration curve for the dose–estimation problem.

Inverse regression

- Dose estimation is an inverse regression problem:
  - We have a model for the dicentric count, y_i, given dose x_i.
  - In practice, we want to estimate x_i given y_i!
- For instance, assume a patient has been admitted to hospital due to potential radiation exposure. A sample of $n_0 = 200$ lymphocytes was analyzed, yielding $y_0 = 150$ dicentrics.



Dose estimation from calibration curve

- Mathematically, this is not a big problem. Assume the observed ratio of dicentrics is $R = \frac{y_0}{n_0}$. Then we have

$$R = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$$

which can be solved wrt x as

$$\hat{x} = \frac{-\hat{\beta}_1 + \sqrt{\hat{\beta}_1^2 - 4\hat{\beta}_2(\hat{\beta}_0 - R)}}{2\hat{\beta}_2}$$

- With $R = \frac{150}{200} = 0.75$, this gives

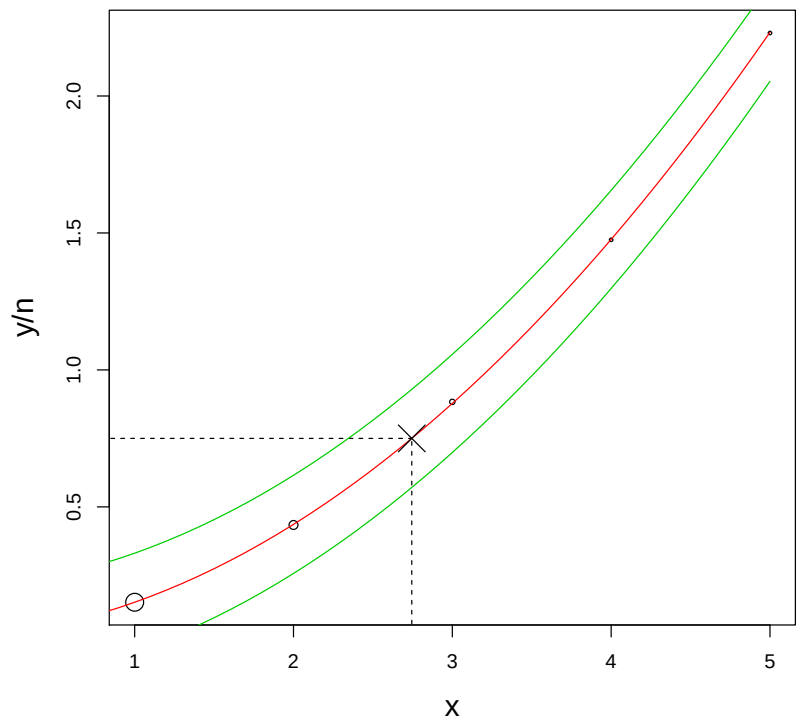
$$\hat{x} = 2.745.$$

Uncertainty

- Of course, this estimation is not *exact*.
- Can one specify the uncertainty in this process?

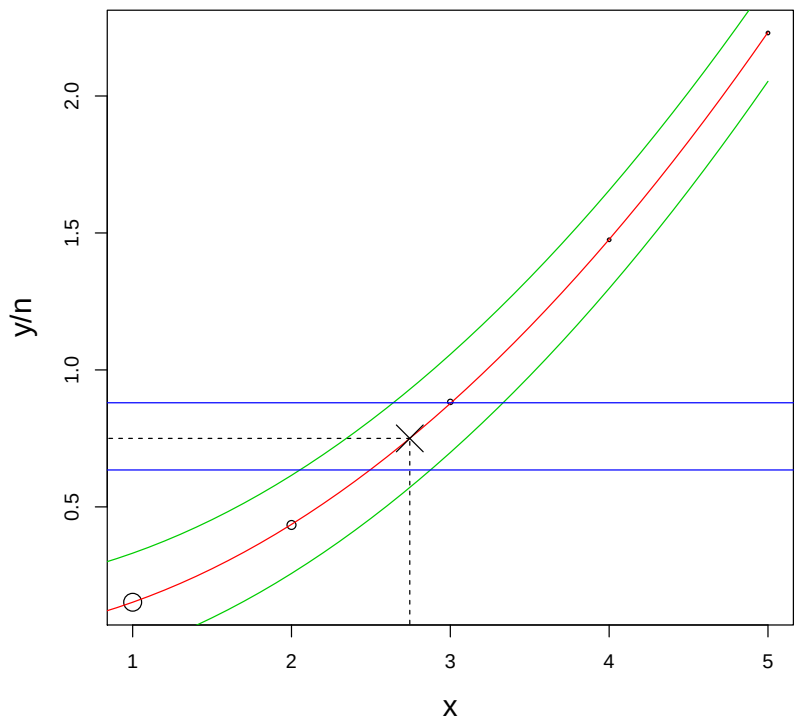
Uncertainty

- Of course, this estimation is not *exact*.
- Can one specify the uncertainty in this process?
- Two sources of uncertainty:
  - Uncertainty due to the estimation of the curve (randomness of the calibration data).



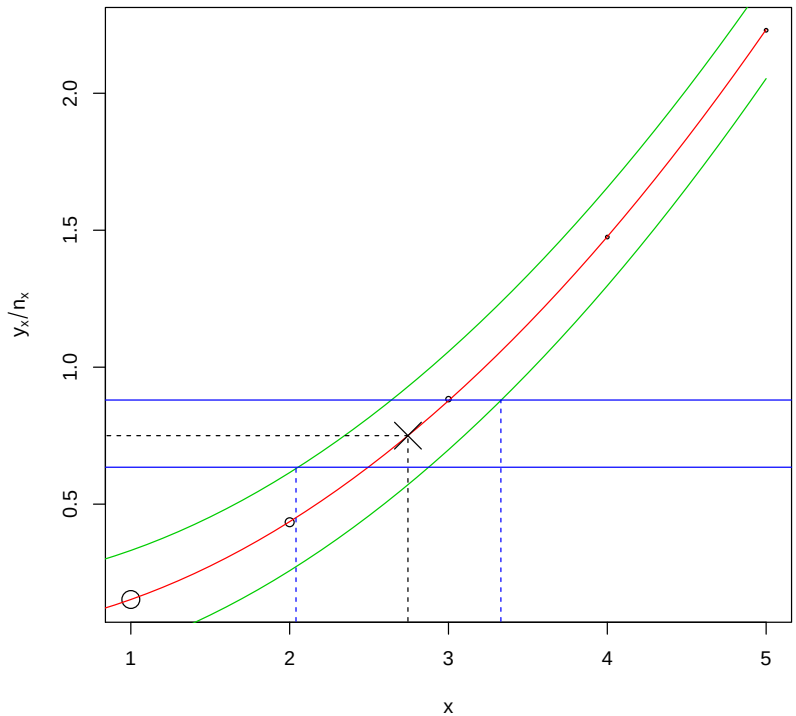
Uncertainty

- Of course, this estimation is not *exact*.
- Can one specify the uncertainty in this process?
- Two sources of uncertainty:
  - Uncertainty due to the estimation of the curve (randomness of the calibration data).
  - Uncertainty due to random variation of y_0, given x_0.



Uncertainty bounds

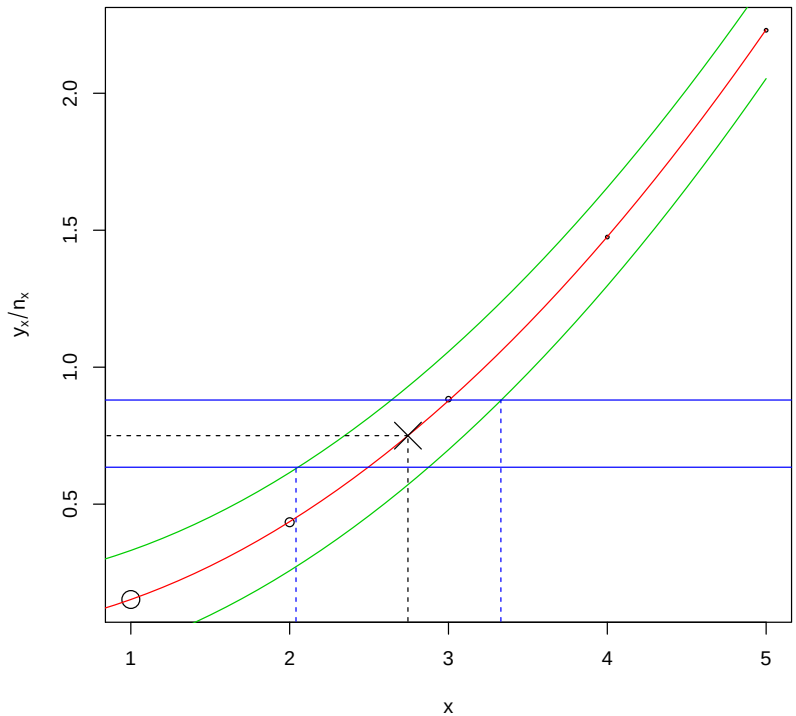
- Combine the two sources of uncertainty ('Merkle's method', 1983):



- Here, a 95% confidence interval for the 'true' dose, x_0, is given as $[2.04, 3.33]$.

Uncertainty bounds

- Combine the two sources of uncertainty ('Merkle's method', 1983):



- Here, a 95% confidence interval for the 'true' dose, x_0, is given as $[2.04, 3.33]$.
- Official uncertainty assessment routine suggested by the International Atomic Energy Agency [IAEA].

A (semi–)Bayesian approach to uncertainty assessment

- Estimate *in vitro* dose-response curve $\hat{\lambda}_i$ as before.

A (semi–)Bayesian approach to uncertainty assessment

- Estimate *in vitro* dose-response curve $\hat{\lambda}_i$ as before.
- For the (potentially) exposed patient, count dicentrics, y_0, in a sample of n_0 cells, yielding 'test data likelihood'

$$L(y_0|\lambda, n_0) \propto e^{-n_0\lambda}\lambda^{y_0}$$

where $\lambda = \beta_0 + \beta_1 x + \beta_2 x^2$, and x the (true, unknown) dose.

A (semi–)Bayesian approach to uncertainty assessment

- Estimate *in vitro* dose-response curve $\hat{\lambda}_i$ as before.
- For the (potentially) exposed patient, count dicentrics, y_0, in a sample of n_0 cells, yielding 'test data likelihood'

$$L(y_0|\lambda, n_0) \propto e^{-n_0\lambda}\lambda^{y_0}$$

  where $\lambda = \beta_0 + \beta_1 x + \beta_2 x^2$, and x the (true, unknown) dose.
- Assume a prior density $p(x, \lambda) = \phi(\lambda|x)p(x)$, where $\phi(\lambda|x) \sim N(\hat{\lambda}, \mathrm{Var}(\hat{\lambda}))$.

A (semi–)Bayesian approach to uncertainty assessment

- Estimate *in vitro* dose-response curve $\hat{\lambda}_i$ as before.
- For the (potentially) exposed patient, count dicentrics, y_0, in a sample of n_0 cells, yielding 'test data likelihood'

$$L(y_0|\lambda, n_0) \propto e^{-n_0\lambda}\lambda^{y_0}$$

  where $\lambda = \beta_0 + \beta_1 x + \beta_2 x^2$, and x the (true, unknown) dose.
- Assume a prior density $p(x, \lambda) = \phi(\lambda|x)p(x)$, where $\phi(\lambda|x) \sim N(\hat{\lambda}, \mathrm{Var}(\hat{\lambda}))$.
- Use Bayes's theorem to obtain posterior density for (λ, x):

$$p(\lambda, x|y_0) \propto L(y_0|\lambda, n_0)p(x, \lambda)$$

A (semi–)Bayesian approach to uncertainty assessment

- Estimate *in vitro* dose-response curve $\hat{\lambda}_i$ as before.
- For the (potentially) exposed patient, count dicentrics, y_0, in a sample of n_0 cells, yielding 'test data likelihood'

$$L(y_0|\lambda, n_0) \propto e^{-n_0\lambda}\lambda^{y_0}$$

  where $\lambda = \beta_0 + \beta_1 x + \beta_2 x^2$, and x the (true, unknown) dose.
- Assume a prior density $p(x, \lambda) = \phi(\lambda|x)p(x)$, where $\phi(\lambda|x) \sim N(\hat{\lambda}, \mathrm{Var}(\hat{\lambda}))$.
- Use Bayes's theorem to obtain posterior density for (λ, x):

$$p(\lambda, x|y_0) \propto L(y_0|\lambda, n_0)p(x, \lambda)$$

- Integration over λ gives *calibrative density of* x:

$$p(x|y_0) \propto p(x) \int L(y_0|\lambda, n_0)\phi(\lambda|x)\, d\lambda$$

A (semi–)Bayesian approach to uncertainty assessment

- Estimate *in vitro* dose-response curve $\hat{\lambda}_i$ as before.
- For the (potentially) exposed patient, count dicentrics, y_0, in a sample of n_0 cells, yielding 'test data likelihood'

$$L(y_0|\lambda, n_0) \propto e^{-n_0\lambda}\lambda^{y_0}$$

  where $\lambda = \beta_0 + \beta_1 x + \beta_2 x^2$, and x the (true, unknown) dose.
- Assume a prior density $p(x, \lambda) = \phi(\lambda|x)p(x)$, where $\phi(\lambda|x) \sim N(\hat{\lambda}, \mathrm{Var}(\hat{\lambda}))$.
- Use Bayes's theorem to obtain posterior density for (λ, x):

$$p(\lambda, x|y_0) \propto L(y_0|\lambda, n_0)p(x, \lambda)$$

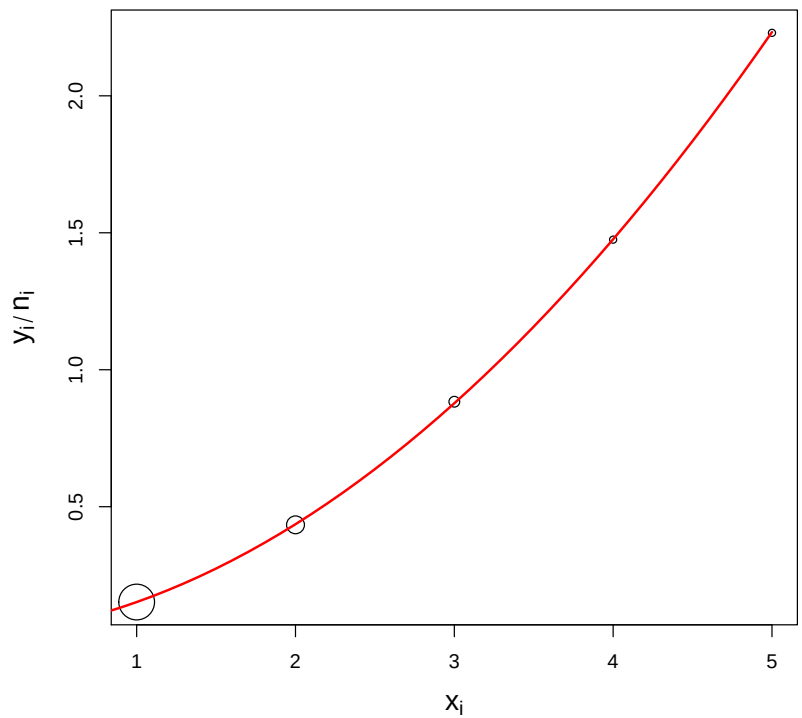
- Integration over λ gives *calibrative density of* x:

$$p(x|y_0) \propto p(x) \int L(y_0|\lambda, n_0)\phi(\lambda|x)\, d\lambda$$

- Integral has explicit solution via Hermite distribution (Higueras et al, 2015)

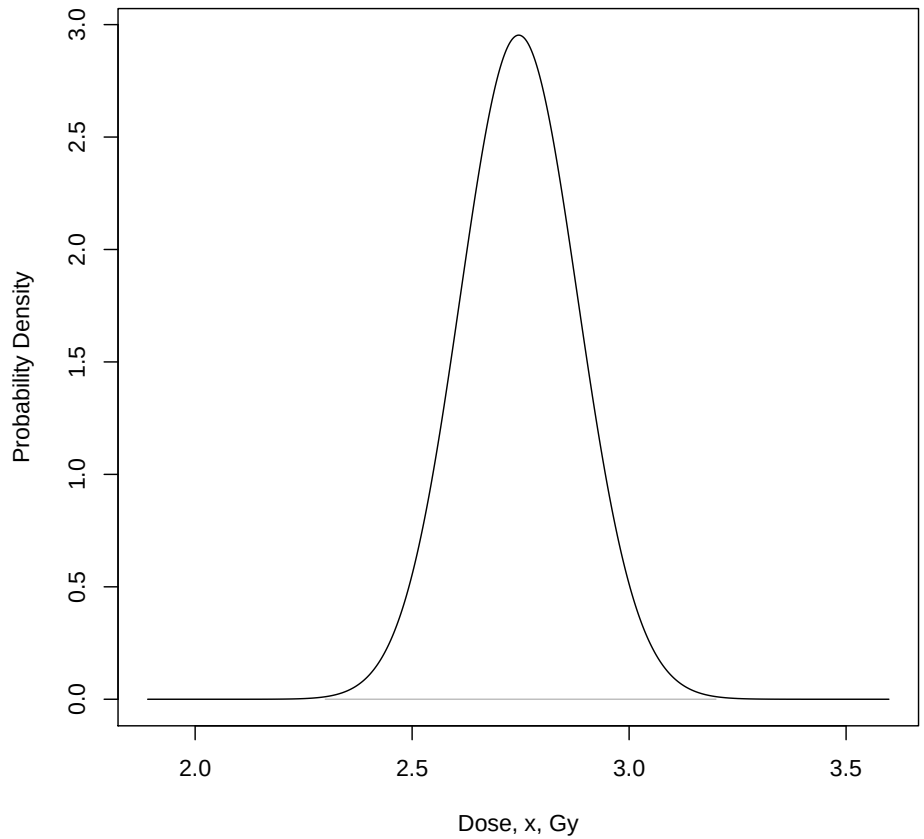
A (semi–)Bayesian approach to uncertainty assessment

- Consider again the example before: Patient sample with $n_0 = 200$, $y_0 = 150$.
- Use the same estimated dose–response curve, $\hat{\lambda}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2$, as before:



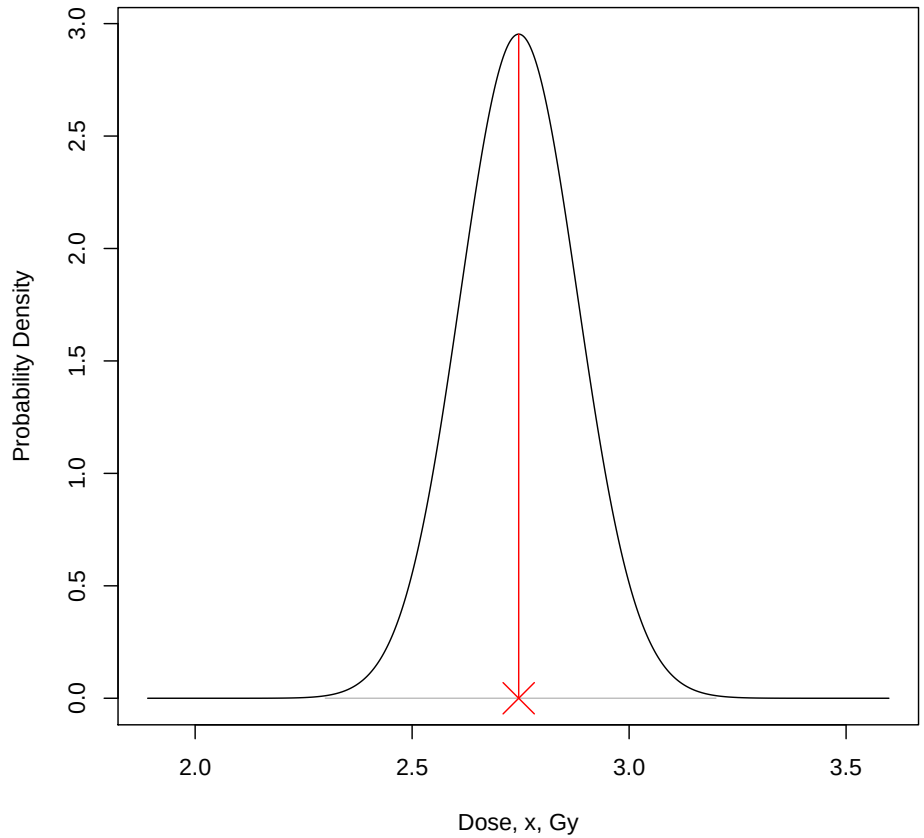
A (semi–)Bayesian approach to uncertainty assessment

- Calibrative density for 'true' dose x, using R package **radir**:



A (semi–)Bayesian approach to uncertainty assessment

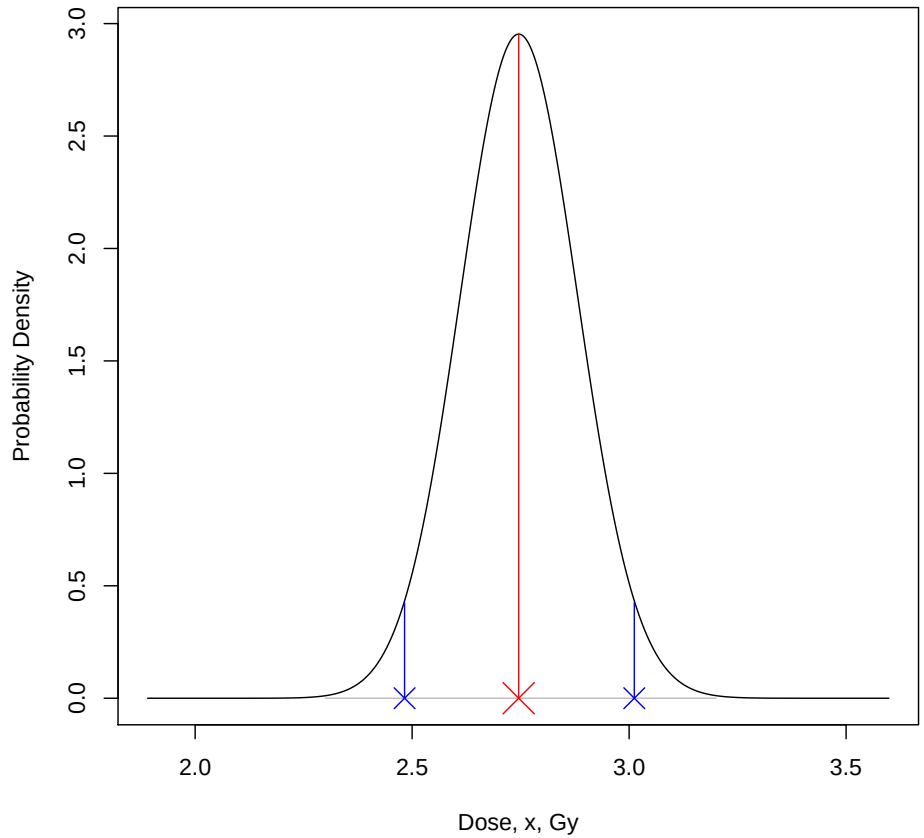
- Dose estimate: Mode of calibrative density:



- $\hat{x} = 2.75$.

A (semi–)Bayesian approach to uncertainty assessment

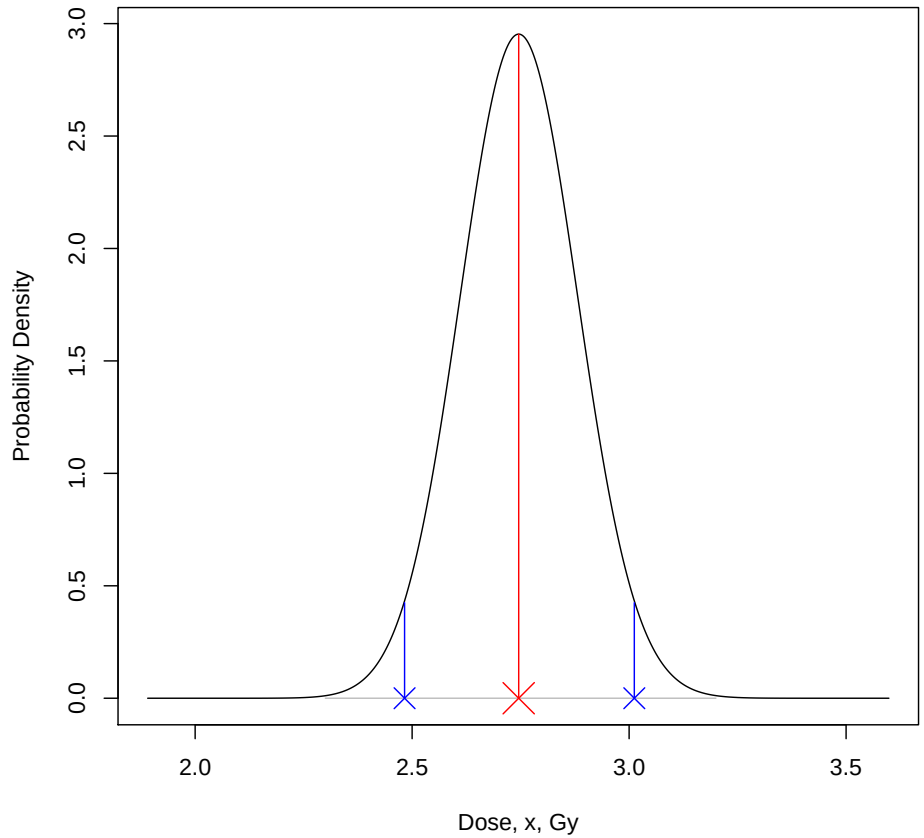
- Uncertainty assessment: 95% Credible intervals



- $CI = [2.48, 3.01]$

A (semi–)Bayesian approach to uncertainty assessment

- Uncertainty assessment: 95% Credible intervals



- $CI = [2.48, 3.01] \in [2.04, 3.33]$

It's not as easy...

- NIHR–funded project (2014): Identification of appropriate response models for chromosomal aberration counts.

It's not as easy...

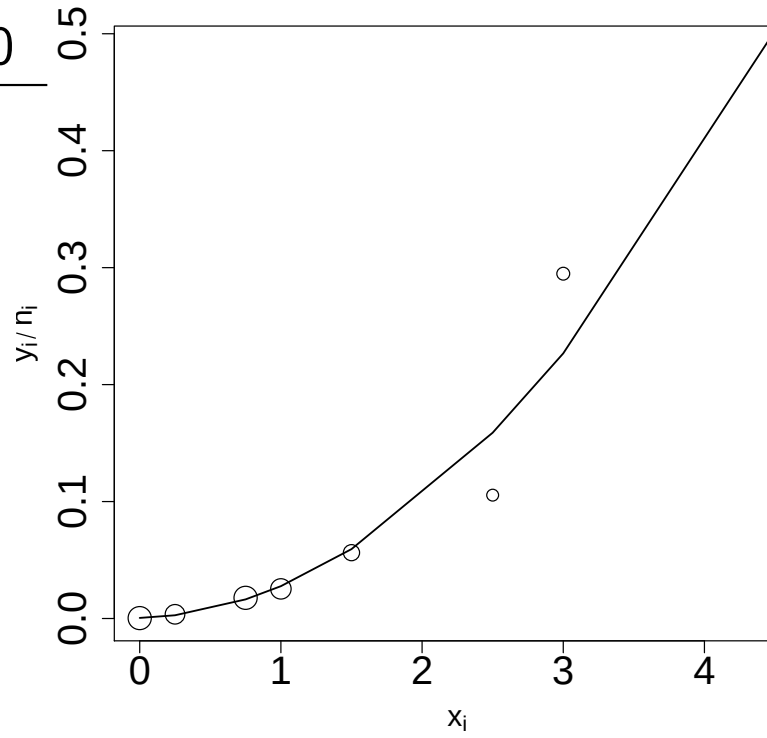
- NIHR–funded project (2014): Identification of appropriate response models for chromosomal aberration counts.
- Initially, carried out an extensive analysis of 11 *in vitro* calibration data sets.

It's not as easy...

- NIHR–funded project (2014): Identification of appropriate response models for chromosomal aberration counts.
- Initially, carried out an extensive analysis of 11 *in vitro* calibration data sets.
- In turned out that the data set shown initially in this talk was the only one for which the Poisson assumption is (approximately) adequate.
- In most occasions, things are not quite as nice...
  - There is overdispersion (variance $>>$ mean)
  - There is zero–inflation (more zero counts than one would expect under the Poisson model)

Example: Cobalt–60 γ– rays

	y_{ij}							
x_i	0	1	2	3	4	5	n_i	y_i
0.00	2591	1	0	0	0	0	2592	1
0.25	2185	8	0	0	0	0	2193	8
0.75	2550	44	1	0	0	0	2595	46
1.00	2231	54	2	0	0	0	2287	58
1.50	1712	96	3	0	0	0	1811	102
2.50	1196	123	7	1	0	0	1327	140
3.00	1070	320	41	6	1	0	1438	424
4.50	895	360	110	25	5	1	1396	680



Alternative models?

- If the response distribution is incorrectly specified, the uncertainty assessment will be incorrect.

Alternative models?

- If the response distribution is incorrectly specified, the uncertainty assessment will be incorrect.
- Alternative models include
  - Zero–inflated regression model (ZIP)

$$\mathrm{P}(Y_{ij} = y_{ij}) = \begin{cases} p + (1-p)\exp(-\lambda_i), & y_{ij} = 0, \\ (1-p)\exp(-\lambda_i)\lambda_i^{y_i}/y_i!, & y_{ij} > 0, \end{cases}$$

  - Negative Binomial model (NB)
  - Zero–inflated Negative Binomial model (ZINB)
  - ...

Alternative models?

- If the response distribution is incorrectly specified, the uncertainty assessment will be incorrect.
- Alternative models include
  - Zero–inflated regression model (ZIP)

$$\mathrm{P}(Y_{ij} = y_{ij}) = \begin{cases} p + (1-p)\exp(-\lambda_i), & y_{ij} = 0, \\ (1-p)\exp(-\lambda_i)\lambda_i^{y_i}/y_i!, & y_{ij} > 0, \end{cases}$$

  - Negative Binomial model (NB)
  - Zero–inflated Negative Binomial model (ZINB)
  - ...
- However, the presence/type of violation of the Poisson model will often not be obvious (from the fitted calibration curve).

Alternative models?

- If the response distribution is incorrectly specified, the uncertainty assessment will be incorrect.
- Alternative models include
  - Zero–inflated regression model (ZIP)

$$\mathrm{P}(Y_{ij} = y_{ij}) = \begin{cases} p + (1-p)\exp(-\lambda_i), & y_{ij} = 0, \\ (1-p)\exp(-\lambda_i)\lambda_i^{y_i}/y_i!, & y_{ij} > 0, \end{cases}$$

  - Negative Binomial model (NB)
  - Zero–inflated Negative Binomial model (ZINB)
  - ...
- However, the presence/type of violation of the Poisson model will often not be obvious (from the fitted calibration curve).
- Tools which can be used to assist in this question:
  - Model selection criteria, e.g AIC, BIC
  - Statistical tests, e.g. LR tests, Score tests
  - Diagnostic plot

Model selection

- Assume we we need to decide between one of the response distributions Poisson, ZIP, NB, ZINB.
- Use Akaike Information Criterion: Find model which minimizes $AIC = -2 \log L + 2p$ (Goodness–of-fit/ complexity trade-off).

Model selection

- Assume we we need to decide between one of the response distributions Poisson, ZIP, NB, ZINB.
- Use Akaike Information Criterion: Find model which minimizes $AIC = -2 \log L + 2p$ (Goodness–of-fit/ complexity trade-off).
- For instance, for the two data sets displayed so far,

AIC for...	X–rays	γ–rays
Poisson	**7622.55**	7504.73
Zero–inflated Poisson (ZIP)	7623.57	7490.36
Negative Binomial (NB)	7624.56	**7489.10**
ZINB	7626.39	7491.44

  - For the X–rays, the Poisson model well supported.
  - For the γ–rays, evidence for NB or perhaps ZIP model.

Model selection

- Assume we we need to decide between one of the response distributions Poisson, ZIP, NB, ZINB.
- Use Akaike Information Criterion: Find model which minimizes $AIC = -2 \log L + 2p$ (Goodness–of-fit/ complexity trade-off).
- For instance, for the two data sets displayed so far,

AIC for...	X–rays	γ–rays
Poisson	**7622.55**	7504.73
Zero–inflated Poisson (ZIP)	7623.57	7490.36
Negative Binomial (NB)	7624.56	**7489.10**
ZINB	7626.39	7491.44

  - For the X–rays, the Poisson model well supported.
  - For the γ–rays, evidence for NB or perhaps ZIP model.
- Model selection methods...
  - give a useful *indication* of a suitable model;
  - are easily implemented;
  - do *not* tell whether one model is *significantly* better than another.

Score test

- Say $H_0 = Po(\lambda_i)$, $H_1 = ZIP(p, \lambda_i)$,
  - in other words, $H_0 : p = 0$.
- Score test statistic:

$$T = S(0, \hat{\beta})^T J(0, \hat{\beta})^{-1} S(0, \hat{\beta})$$

  where $S = \partial L / \partial \beta$ and $J = \partial L / \partial \beta \beta^T$ are the score function and Fisher information of model H_1, evaluated under the model fit β under H_0.
- Critical value for $\alpha = 0.05$ is $\chi^2_{1,0.95} = 3.84$.

Score test

- Say $H_0 = Po(\lambda_i)$, $H_1 = ZIP(p, \lambda_i)$,
  - in other words, $H_0 : p = 0$.
- Score test statistic:

$$T = S(0, \hat{\beta})^T J(0, \hat{\beta})^{-1} S(0, \hat{\beta})$$

  where $S = \partial L / \partial \beta$ and $J = \partial L / \partial \beta \beta^T$ are the score function and Fisher information of model H_1, evaluated under the model fit β under H_0.
- Critical value for $\alpha = 0.05$ is $\chi^2_{1,0.95} = 3.84$.
- Results for 8 data sets (Oliveira et al, 2016):

	Homogenous exposure				Partial exposure			
LET	low		high		low		high	
id	0.92	18.17	87.72	61.32	2007.39	1418.28	416.20	387.91

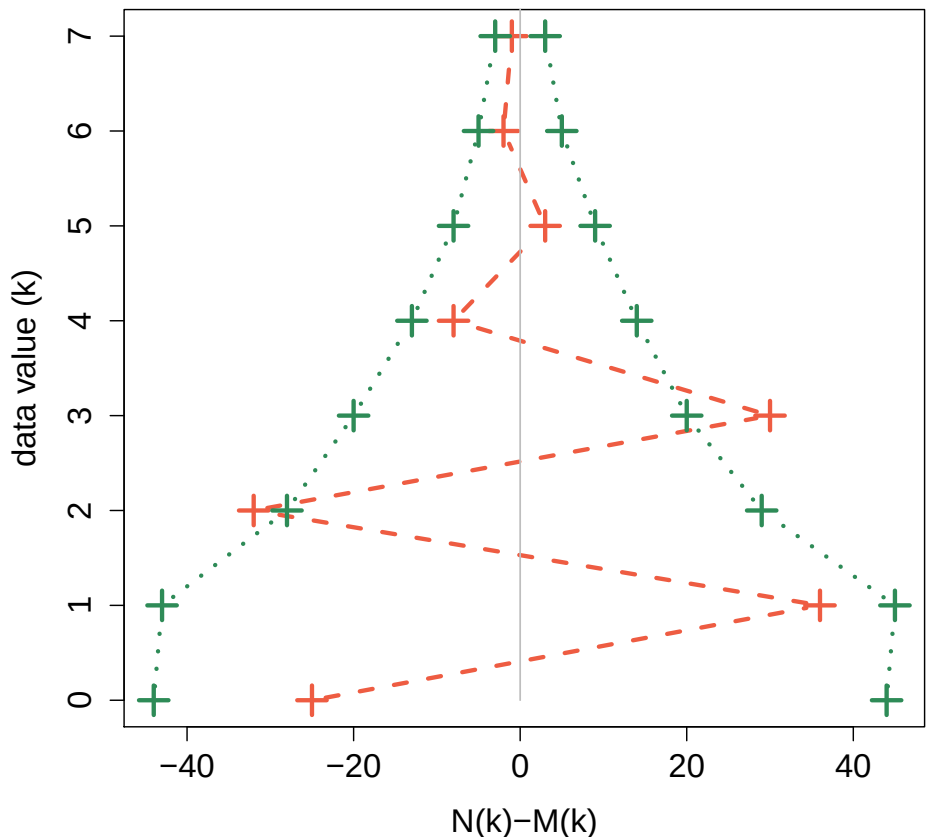
  - (the first two ones are the data sets discussed previously)
  - All data sets except the first one (X–rays) are zero–inflated!!

New graphical device

- We developed a new graphical tool to detect zero–inflation (and in fact, any–number–inflation).
- Effectively, it is demonstrated whether the number of counts, for $k = 0, 1, 2, ...,$ is consistent with the specified count distribution.

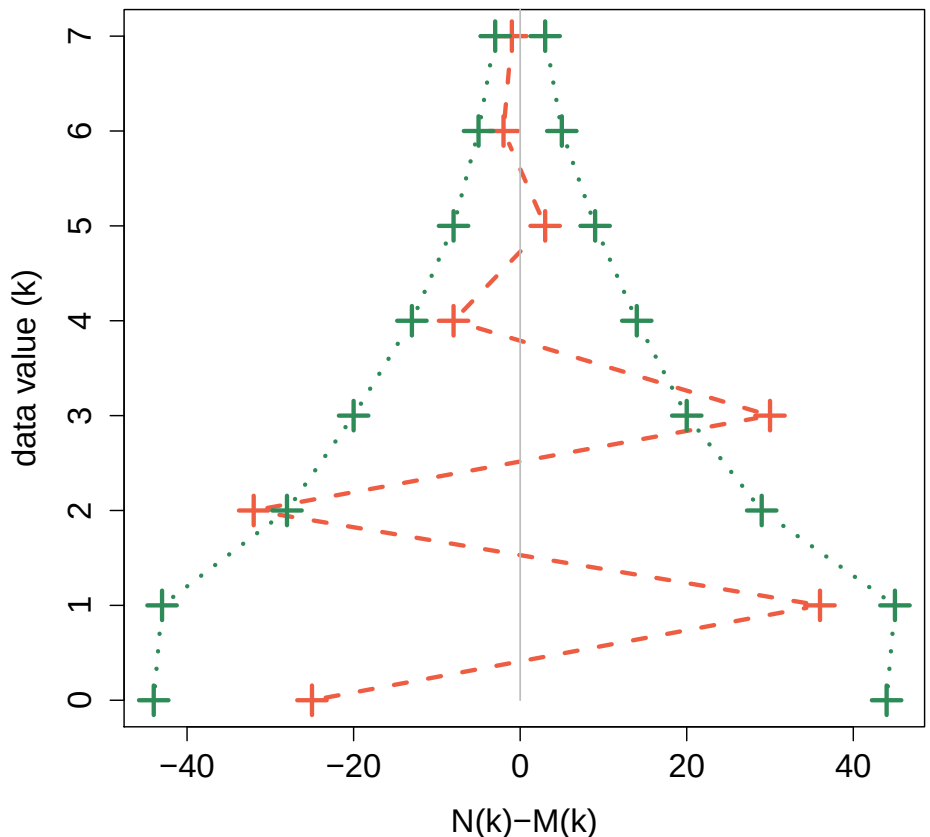
New graphical device

- We developed a new graphical tool to detect zero–inflation (and in fact, any–number–inflation).
- Effectively, it is demonstrated whether the number of counts, for $k = 0, 1, 2, ...,$ is consistent with the specified count distribution.
- For the X–ray calibration data (first data set),



New graphical device

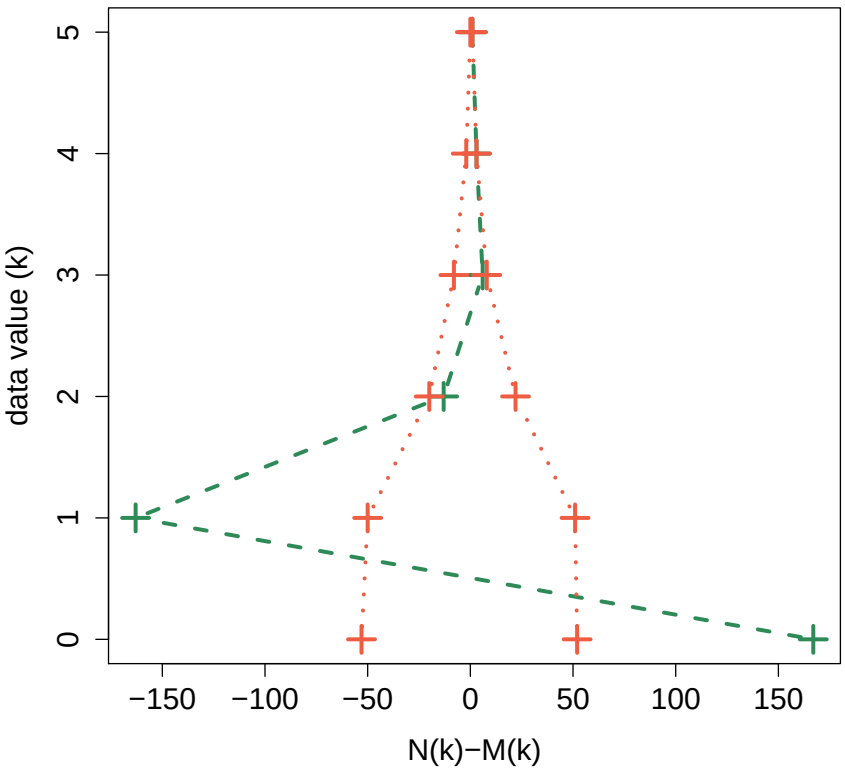
- We developed a new graphical tool to detect zero–inflation (and in fact, any–number–inflation).
- Effectively, it is demonstrated whether the number of counts, for $k = 0, 1, 2, ...$, is consistent with the specified count distribution.
- For the X–ray calibration data (first data set),



'Christmas tree plot'
(Einbeck & Wilson, 2016)

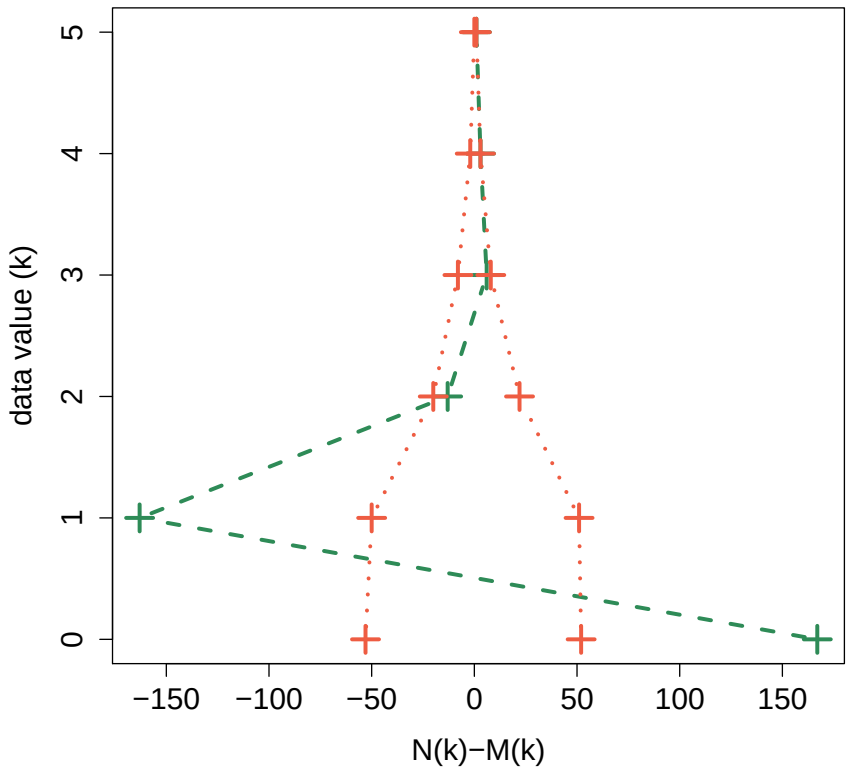
Christmas tree plot for γ–ray data

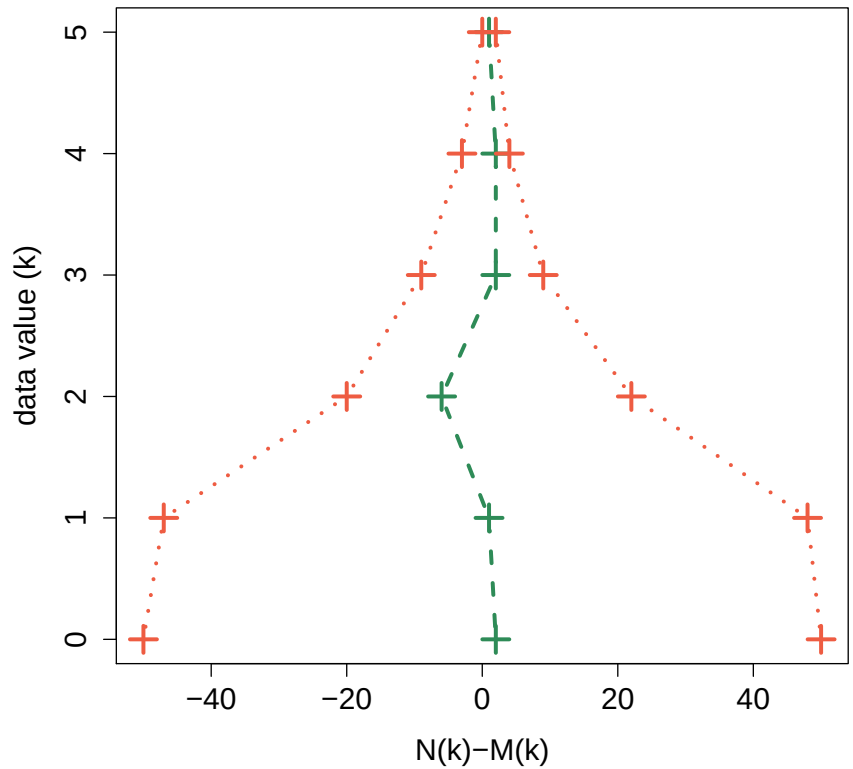
- using Poisson model



Christmas tree plot for γ–ray data

- using Poisson model
- using ZIP model





Summary: Model choices

- It can be concluded (Oliveira et al, 2016):
  - Zero–inflation is driven by partial (body) exposure;
  - Overdispersion is driven by densely ionizing radiation, *i.e.* radiation with high linear energy transfer (LET);

Summary: Model choices

- It can be concluded (Oliveira et al, 2016):
  - Zero–inflation is driven by partial (body) exposure;
  - Overdispersion is driven by densely ionizing radiation, *i.e.* radiation with high linear energy transfer (LET);
  - the absence of partial exposure/ densely ionizing radiation does not guarantee the absence of zero–inflation/overdispersion.
- Recommended model choices

exposure		whole body	partial
LET	low	Poisson/NB	ZIP
	high	NB/Neyman A	ZINB

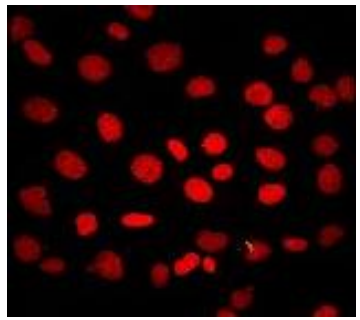
- Some work has been done to extend dose (uncertainty) estimation routines to NB/ZIP models; further research required (Higueras et al, 2015)

Strength and limitations of dicentric assay

- Strength:
  - very little inter–individual (and inter–laboratory) variation;
  - dicentric chromosome aberrations can be identified still several weeks after exposure;
  - internationally accepted ‘gold–standard’.
- Limitations:
  - needs 3–4 days after irradiation before dicentric chromosomes become visible (need to reach metaphase);
  - relatively work–intensive and expensive methodology;
  - not viable for large–scale radiation accidents.
- Hence, alternative biomarkers have recently been considered.

γ– H2AX data

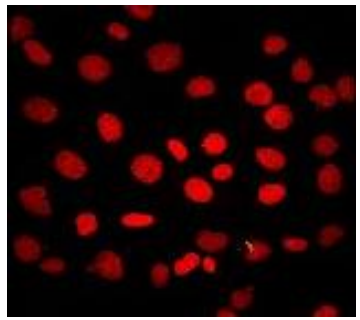
- Relatively new technology: Protein biomarker
- Double strand breaks lead to 'phosphorylation' of the H2AX protein, yielding γ–H2AX foci.
- γ–H2AX foci are counted using flow cytometers or immunofluorescence microscopy.



- Quicker, cheaper, less invasive than dicentric array.

γ– H2AX data

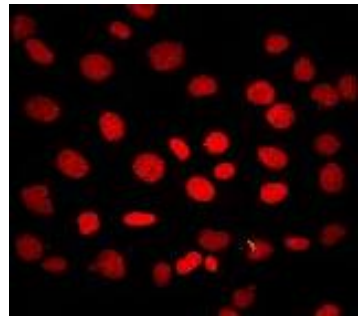
- Relatively new technology: Protein biomarker
- Double strand breaks lead to 'phosphorylation' of the H2AX protein, yielding γ–H2AX foci.
- γ–H2AX foci are counted using flow cytometers or immunofluorescence microscopy.



- Quicker, cheaper, less invasive than dicentric array.
- Problems:
  - only visible up to 24 hours after exposure;
  - large inter–individual variation, hence requires further calibration steps, or models which can capture this variation.

γ– H2AX data

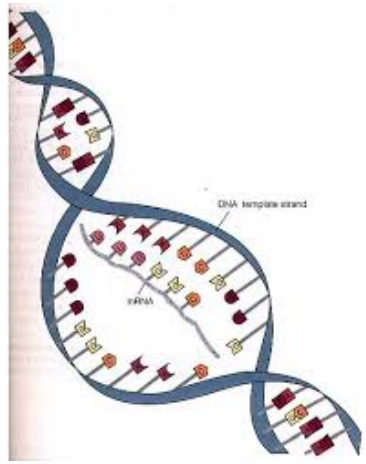
- Relatively new technology: Protein biomarker
- Double strand breaks lead to 'phosphorylation' of the H2AX protein, yielding γ–H2AX foci.
- γ–H2AX foci are counted using flow cytometers or immunofluorescence microscopy.



- Quicker, cheaper, less invasive than dicentric array.
- Problems:
  - only visible up to 24 hours after exposure;
  - large inter–individual variation, hence requires further calibration steps, or models which can capture this variation.
- Some ad–hoc dose estimation methods available (Ainsbury et al, 2016).
- Critical gap remains: Fast ($< 24h$) dose assessment with samples that taken $> 24h$ after the radiation incident.

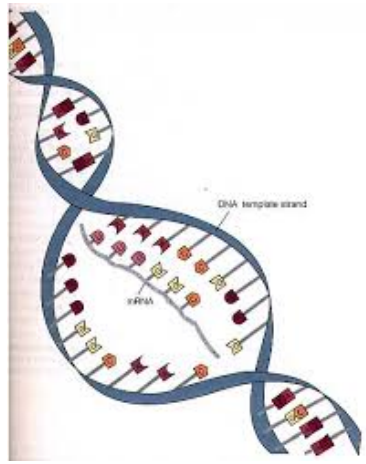
Microarray–based biomarkers

- Recently, it has been demonstrated that certain genes respond to ionizing radiation with a change in their gene expression level.
- Genes are expressed by production of mRNAs from DNA, and protein from mRNAs.
- mRNA is a linear molecule which carries a copy of the gene to be expressed from the nucleus.

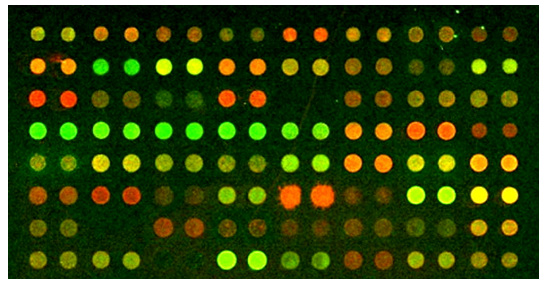


Microarray–based biomarkers

- Recently, it has been demonstrated that certain genes respond to ionizing radiation with a change in their gene expression level.
- Genes are expressed by production of mRNAs from DNA, and protein from mRNAs.
- mRNA is a linear molecule which carries a copy of the gene to be expressed from the nucleus.

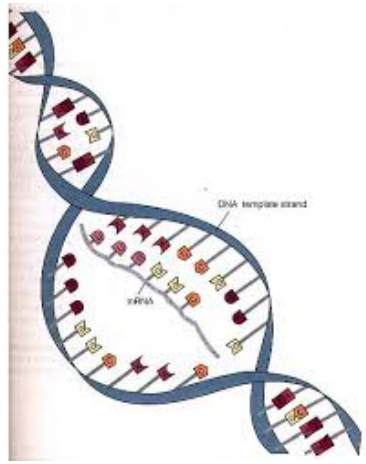


- Expressions can be measured using
  - Hybridization based methods (e.g. microarrays, usually continuous data);
  - Sequencing techniques (usually count data).

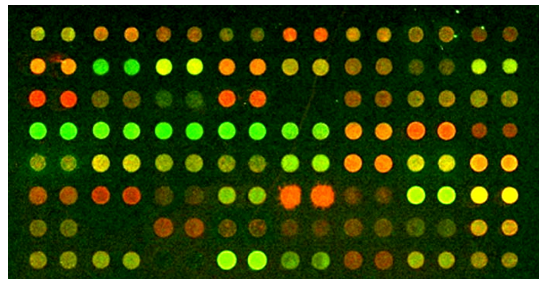


Microarray–based biomarkers

- Recently, it has been demonstrated that certain genes respond to ionizing radiation with a change in their gene expression level.
- Genes are expressed by production of mRNAs from DNA, and protein from mRNAs.
- mRNA is a linear molecule which carries a copy of the gene to be expressed from the nucleus.

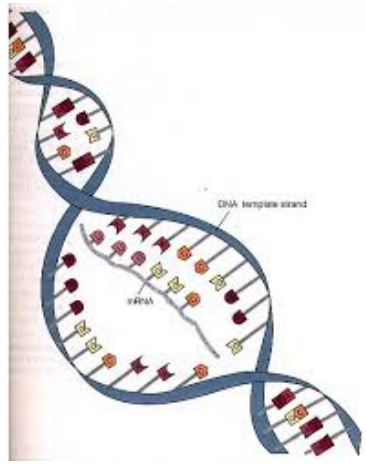


- Expressions can be measured using
  - Hybridization based methods (e.g. microarrays, usually continuous data);
  - Sequencing techniques (usually count data).

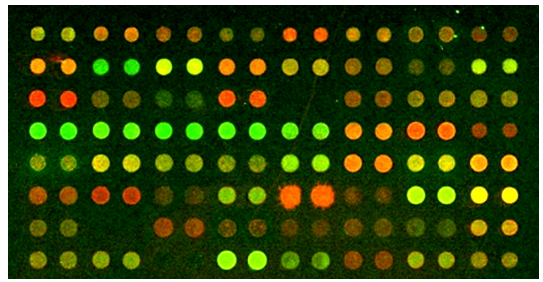


Microarray–based biomarkers

- Recently, it has been demonstrated that certain genes respond to ionizing radiation with a change in their gene expression level.
- Genes are expressed by production of mRNAs from DNA, and protein from mRNAs.
- mRNA is a linear molecule which carries a copy of the gene to be expressed from the nucleus.



- Expressions can be measured using
  - Hybridization based methods (e.g. microarrays, usually continuous data);
  - Sequencing techniques (usually count data).



- Several radiation–responsive genes are known.
- Again, substantial inter–individual variation.
- No reliable statistical methodology for dose estimation yet.

References

(1) Merkle W (1983). Statistical methods in regression and calibration analysis of chromosome aberration data. *Radiation and Environmental Biophysics* **21**, 217–233.

(2) Higueras M, Puig P, Ainsbury EA & Rothkamm K (2015). A new inverse regression model applied to radiation biodosimetry. *Proc's of the Royal Society* **471**, DOI: 10.1098/rspa.2014.0588.

(3) 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. *Biometrical Journal* **58**, 259–279.

(4) Ainsbury EA, Higueras M, Puig P, Einbeck J et al. (2016), Uncertainty of fast biological radiation dose assessment for emergency response scenarios. *International Journal of Radiation Biology*, DOI: 10.1080/09553002.2016.1227106

(5) Einbeck J & Wilson P (2016). A diagnostic plot for assessing model fit in count data models. In: Dupuys J-F & Josse J (Eds). Proc's of the 31st IWSM, Rennes, France, 4-8/7/2016, 103–108.