DoseMatic – the guide to the statistical methods

Maciej Lewicki

October 4, 2017

Based on:
INTERNATIONAL ATOMIC ENERGY AGENCY, VIENNA, 2001
"Cytogenetic Analysis for Radiation Dose Assessment – A Manual"
`http://www-pub.iaea.org/MTCD/publications/PDF/TRS405_scr.pdf`

...But a little more explicit.

Contents

Part I

Acute Exposure Dose Estimation

1 Dose estimation uncertainty calculation

In order to express the uncertainty of the dose assessment the 95% confidence interval (later refered as CI95%) is chosen as a reasonable limit. It defines an interval of 95% probability of enclosing the true dose.
The confidence limits are affected by two sources of uncertainty:

- Poisson nature of of the yields aberrations
- and uncertainties of the calibration curve parameters (following normal distribution).

The paper refers to three different approaches, none of them being the exact one, but each has a region of usability.
At first let us derive the basic statistics associated with the calibration curve:

- The calibration curve equation:
$$Y(D) = c + \alpha D + \beta D^2 \tag{1}$$
- The inverse function, extracting the dose D:
$$D(Y) = \frac{2}{\beta}\left(-\alpha + \sqrt{\alpha^2 + 4\beta(Y-c)}\right) \tag{2}$$
- The calibration curve (eq.1) differentials with respect to fitting parameters:
$$\frac{\partial Y}{\partial c} = 1, \quad \frac{\partial Y}{\partial}\alpha = D, \quad \frac{\partial Y}{\partial \beta} = D^2 \tag{3}$$
- The uncertainty of calibration curve:
$$\begin{aligned} u_{\text{fit}}(Y) &= \frac{\partial Y}{\partial c}\cdot u(c) + \frac{\partial Y}{\partial \alpha}\cdot u(\alpha) + \frac{\partial Y}{\partial \beta}\cdot u(\beta) \\ u_{\text{fit}}(Y) &= \quad u(c) \quad + \quad D\cdot u(\alpha) \quad + \quad D^2\cdot u(\beta) \end{aligned} \tag{4}$$
- The inverse function (eq.2) differentials with respect to fitting parameters:
$$\frac{\partial D}{\partial c} = \frac{4}{\sqrt{\alpha^2 + 4\beta(Y-c)}} \tag{5a}$$
$$\frac{\partial D}{\partial \alpha} = -\frac{2}{\beta} + \frac{2\alpha}{\beta\sqrt{\alpha^2 + 4\beta(Y-c)}} \tag{5b}$$
$$\frac{\partial D}{\partial \beta} = \frac{4(Y-c)}{\beta\sqrt{\alpha^2 + 4\beta(Y-c)}} - \frac{2\left(\sqrt{\alpha^2 + 4\beta(Y-c)} - a\right)}{\beta^2} \tag{5c}$$
- Variance of arbitrary variable X:
$$\text{var}(X) = E\left[(X-\bar{X})^2\right] = \frac{1}{n}\sum_i^n (X_i - \bar{X})^2 \tag{6}$$

- Covariance of arbitrary variables X and Y:

$$\operatorname{cov}(X) = E\left[(X-\bar{X})(Y-\bar{Y})\right] = \frac{1}{n}\sum_{i}^{n}(X_i-\bar{X})(Y_i-\bar{Y}) \tag{7}$$

- Standard error of arbitrary variable X:

$$\sigma(X) = \sqrt{\operatorname{var}(X)} = \sqrt{\frac{1}{n}\sum_{i}^{n}(X_i-\bar{X})^2} \tag{8}$$

- Confidence intervals of 95% (CI95%) for given *sigma* (eq.8):

$$\text{lower CI95\%(X)} = X - 1.96\cdot\sigma(X) \tag{9a}$$
$$\text{upper CI95\%(X)} = X + 1.96\cdot\sigma(X) \tag{9b}$$

1.1 Method A

There are three steps of the procedure:

1. Calculate the dose D from the inverse curve equation (eq.2) for a measured yield Y.
2. Calculate the variance of dose – $\operatorname{var}(D)$ – from a given equation:

$$\begin{aligned}\operatorname{var}(X) \quad = \quad & \left(\frac{\partial D}{\partial c}\right)^2\cdot\operatorname{var}(c) \;+\; \left(\frac{\partial D}{\partial \alpha}\right)^2\cdot\operatorname{var}(\alpha) \;+\; \left(\frac{\partial D}{\partial \beta}\right)^2\cdot\operatorname{var}(\beta) \;+\\ + \; & \frac{\partial D}{\partial \alpha}\frac{\partial D}{\partial \beta}\cdot\operatorname{cov}(\alpha,\beta) \;+\; \frac{\partial D}{\partial c}\frac{\partial D}{\partial \alpha}\cdot\operatorname{cov}(c,\alpha) \;+\; \frac{\partial D}{\partial c}\frac{\partial D}{\partial \beta}\cdot\operatorname{cov}(c,\beta)\end{aligned} \tag{10}$$

   Where all the derivatives are defined above, the variances of curve parameters are givevn by the fitting routine and the variance of the yield Y is derived on the assumption of a Poisson distribution (pmf – probability mass function, k – the number of counts, n – number of cells, λ – the mean number of counts):

$$Y \quad \equiv \quad \frac{1}{n}\sum_{i}^{n} Y_i \quad \equiv \quad \lambda \tag{11}$$

$$\operatorname{pmf}(Y_i) = P(k,\lambda) = \frac{\lambda^k e^{-k}}{k!} \tag{12}$$

   Thus the variance and σ are equal respectively:

$$\operatorname{var}(Y) = \lambda, \qquad \sigma(Y) = \sqrt{\lambda} \tag{13}$$

3. Now let us derive the CI95%:

$$\begin{aligned}\text{lower CI95\%}(D) = D_L = D - 1.96\cdot\sigma(D)\\ \text{upper CI95\%}(D) = D_U = D + 1.96\cdot\sigma(D)\end{aligned} \tag{14}$$

1.2 Method B

1. Calculate the dose D from the inverse curve equation (eq.2) for a measured yield Y.
2. Estimate Y error using curve's uncertaintiy (eq.4):

$$Y = Y \pm u_{\text{fit}}(Y) \tag{15}$$

3. From the observed yield estimate its Poisson standard error (eq.13):

$$\sigma(Y) = \sqrt{\lambda} \tag{16}$$

4. Add the errors from points 2 and 3 (eq. 15 and 16):

$$u(Y) = \sqrt{[u_{\text{fit}}(Y)]^2 + \lambda} \tag{17}$$

$$Y_L = Y - u(Y), \qquad Y_U = Y + u(Y) \tag{18}$$

5. Now, using the inverse function (eq.2) calculate D_U and D_L:

$$D_L = D(Y_L), \qquad D_U = D(Y_U) \tag{19}$$

1.3 Method C

1. Assuming Poisson distribution, calculate the CI95% on the observed yield:

$$\begin{aligned} Y_L &= Y - 1.96 \cdot \sigma(Y) \\ Y_U &= Y + 1.96 \cdot \sigma(Y) \end{aligned} \tag{20}$$

2. Now determine the intersection of the confidence limits Y_L, Y_U with the calibration curve's upper limit and lower limit respectively (or with the curve itself for simplicity):

$$\begin{aligned} Y_L &= Y(D_L) + u(Y) \quad \rightarrow \quad D_L \\ Y_U &= Y(D_U) - u(Y) \quad \rightarrow \quad D_U \end{aligned} \tag{21}$$

$$\begin{aligned} D_L &= D(Y_L - u(Y)) \\ D_U &= D(Y_U + u(Y)) \end{aligned} \tag{22}$$

Or the simplified, straightforward version:

$$\begin{aligned} D_L &= D(Y_L) \\ D_U &= D(Y_U) \end{aligned} \tag{23}$$

2 Conclusions

Method A:

- Obviously most accurate. Even almost exact – if it were not for the usually wrong assumption on the normal distribution of the aberrations. This is a problem for low number of aberrations, which means low doses.
- This method is recommended for large number of scored aberrations (or high doses).

Method B:

- Suffers from similar defect as above.
- The region of recommended usage is when the uncertainty of measured yields is similar to the uncertainty of the calibration curve.

Method C:

- Takes into account the Poisson nature of the number of aberrations distribution.

- In the basic form overestimates the effect of calibration curve uncertainties.
- In the simplified form the curve uncertainties are not taken into account at all.
- To be used when the uncertainty of yields dominates over the uncertainty of the calibration curve.

Part II

Partial Exposure Dose Estimation

3 Calculation of output values

3.1 Basic Statistics

- Standard deviation of a sample:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

- Standard error – based on standard deviation of a sample:

$$SE = \frac{\sigma}{n} \approx \frac{s}{n}$$

- Dispertion index:

$$D = \frac{\sigma^2}{\mu} \approx \frac{s^2}{\bar{x}}$$

- U-test:

$$u = (D-1) \cdot \sqrt{2 \cdot \left(1 - \frac{1}{\bar{x}}\right)}$$

3.2 Dolphin Method

- Calculation of yield using Dolphin method:

$$\frac{Y}{1-e^{-Y}} = \frac{X}{N-n_0}$$

$$Yf = \frac{X}{N}$$

where:
N – number of cells scored,
X – number of observed dicentrics,
n_0 – cells free of dicentrics,
f – fraction of irradiated body,
Y – yield:

$$Y = W\left(\frac{e^{\frac{X}{n_0-N}}\ X}{n_0-N}\right) - \frac{X}{n_0-N}$$

where $W(\cdots)$ is the principal solution of a Lambert W-function.

3.3 Qdr Method

- Calculation of yield using Qdr method:

$$\mathrm{Qdr} = \frac{X}{N_u} = \frac{Y}{1 - e^{-Y_1 - Y_2}}$$

where:
X – number of dicentrics plus rings,
N_u – number of damaged cells,
Y_1 – yield of dicentrics plus rings,
Y_2 – yield of acentrics,
Qdr – yield of dicentrics and rings among damaged cells,
Y - yield:

$$Y = \mathrm{Qdr} \cdot \left(1 - e^{-Y_1 - Y_2}\right)$$