Non-Clinical Statistics Conference
24-26 September 2012, Potsdam, Germany

Normalizing to control measurements in dose-response studies: What is the effect on parameter estimates?

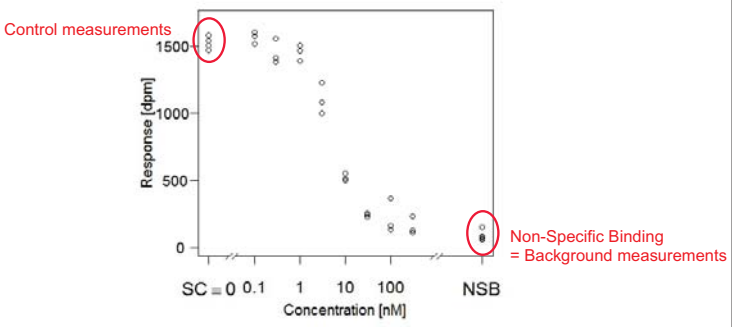
Annette Kopp-Schneider, Xiaoqi Jiang, Oriana Ponta, Sven Stanzel, Alexius Freyberger(*), Marc Weimer

Division of Biostatistics, German Cancer Research Center, Heidelberg
(*) Bayer Pharma AG, Wuppertal, Germany

dkfz. GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION

Dose-response data: Recombinant androgen receptor binding assay

- Experiment with Solvent Control (6 replicates), 6 concentrations (3 rep each) of test compound and Non-Specific Binding (6 rep)
- Response measured in 'disintegrations per minute'



Background correction and normalization

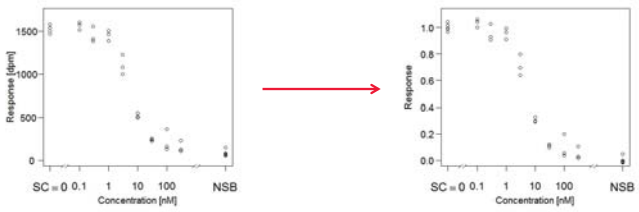
- Motivation:
  Experiments differ in response level
  'Harmonize' graphical display
  e.g. from GraphPad Prism 'Analyzing dose-response data':

> **Normalizing the Y Values**
> In this section, we'll convert the Y values in data sets A and B to a common scale. This is useful when you want to compare the shape or position (EC_{50}) of two or more curves and don't want to be distracted by different maximum and minimum values.

- Background correction (bc):
  Subtract mean(NSB) from all measurements $y_i' = y_i - \overline{NSB}$
- Normalize to control:
  Divide (bc) measurements by mean(bc Control)

$$y_i'' = \frac{y_i - \overline{NSB}}{Control - \overline{NSB}}$$

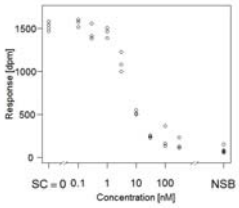
Background correction and normalization



Background correction and normalization: Statistical issues (1)

- Problem: Control values are random
  → Distribution of ratios of normal variables?

- Two situations:
  1. Divide response values for concentration > 0 by mean of control
  2. Divide control values by mean of control



Ratio of normal variables:

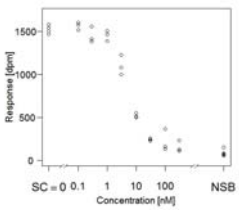
D. Hinkley (1969), On the ratio of two correlated normal variables. Biometrika 56, 635-639

1. Let $X \sim N\left(\mu_X, \sigma_X^2\right)$ and $Y \sim N\left(\mu_Y, \sigma_Y^2\right)$ be independent and $\frac{\sigma_Y}{\mu_Y} \to 0$

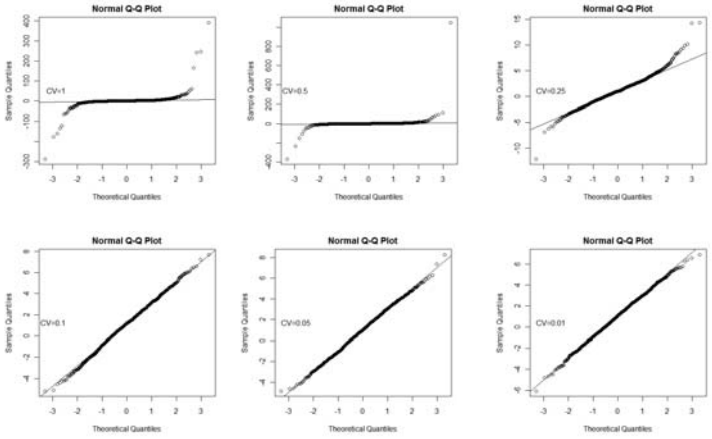
Then asymptotically $\frac{X}{Y} \sim N\left(\frac{\mu_X}{\mu_Y}, \left(\frac{\sigma_X}{\mu_Y}\right)^2\right)$

2. $X_i \sim N\left(\mu_X, \sigma_X^2\right), i = 1,..,n$

Then asymptotically $\frac{X_i}{\bar{X}} \sim N\left(1, \left(\frac{\sigma_X}{\mu_X}\right)^2\right)$

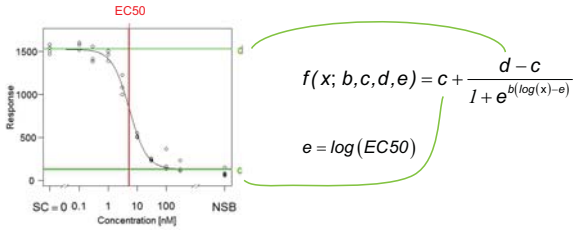


Distribution of X/Y for different Coefficients of Variation of Y



Background correction and normalization: Statistical issues (2)

- Non-linear functions are fitted to dose-response data, e.g. 4-parameter log-logistic function
- Aim: derive summary characteristics, e.g. EC50



What is the impact of data transformation on parameter estimates?

Background correction and normalization in practise

- Transformation $y \mapsto y'' = \alpha(y-\beta),\ \beta = \overline{NSB},\ \alpha = 1/\overline{Control - NSB}$
- Implication for parameters of 4-par log-logistic model:

$$\alpha\left(f(x;b,c,d,e)-\beta\right) = \alpha\left(c + \frac{d-c}{1+exp\left\{b\left[log(x)-e\right]\right\}} - \beta\right)$$

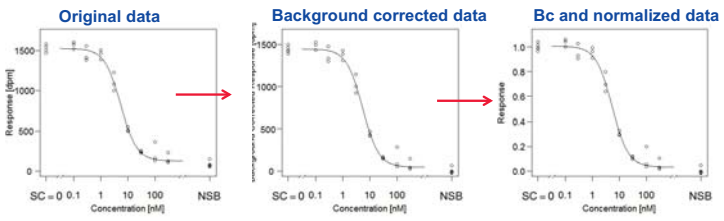
$$= \alpha(c-\beta) + \frac{\alpha(d-\beta)-\alpha(c-\beta)}{1+exp\left\{b\left[log(x)-e\right]\right\}} = f\left(x;b,\alpha(c-\beta),\alpha(d-\beta),e\right).$$

- Residuals:
$y'' - \alpha\left[f(x;...)-\beta\right] = \alpha(y-\beta) - \alpha\left[f(x;...)-\beta\right] = \alpha\left[y - f(x;...)\right]$

Background correction and normalization in practise

Transformation	Hill slope, b	Lower Boundary, c	Upper Boundary, d	log(EC50) e	EC50 [nM]
-	1.45 [1.16,1.75]	127.51 [84,171]	1526.95 [1484,1570]	1.68 [1.53,1.84]	5.39 [4.60, 6.31]
y → α(y-β)	1.45 [1.16,1.75]	0.03 [0.00,0.06]	1.01 [0.98, 1.04]	1.68 [1.53,1.84]	5.39 [4.60, 6.31]

Background correction and normalization in practise



Typical approach by toxicological user

- Use background correction
  → fix c ≡ 0, use 3-par log-logistic model f(x;b,0,d,e)
- Use (background correction and) normalization
  → fix c ≡ 0, d ≡ 1, use 2-par log-logistic model f(x;b,0,1,e)

e.g. from GraphPad Prism 'Analyzing dose-response data':

Constraining Curve-Fit Parameters

Since we normalized the original data such that the vertical range extends by definition from 0 to 100, it doesn't make sense to fit the "bottom" and the "top" of the curves. So we'll fix those parameters, leaving only the midpoint (log EC_{50}) and slope (Hill slope) of each curve to be fitted by Prism. Select the **Constaints** tab. Constrain the parameters BOTTOM and TOP to 0 and 100, respectively:

Impact of fixing model parameters

Fixed parameters	Hill slope, b	Lower Boundary, c	Upper Boundary, d	log(EC50) e	EC50 [nM]
-	1.45 [1.16,1.75]	0.03 [0.00,0.06]	1.01 [0.98, 1.04]	1.68 [1.53,1.84]	5.39 [4.60, 6.31]
c=0	1.32 [1.05,1.59]	-	1.01 [0.98, 1.05]	1.75 [1.58,1.92]	5.76 [4.86, 6.82]
c=0, d=1	1.35 [1.09,1.60]	-	-	1.77 [1.62,1.92]	5.88 [5.06, 6.84]

Conclusions

- For evaluation of single experiment:
  Background correction and normalization
  does not change parameter estimates
  if all 4 parameters of log-logistic model are estimated.
  changes parameter estimates
  if parameters c and/or d are fixed.

- If several experiments are background corrected and normalized separately and experiments are evaluated jointly, then estimates change.

- Model-based transformation using $\hat{c}$ and $\hat{d}$ from fit of 4-par log-logistic model:

$$y \mapsto y'' = \frac{y - \hat{c}}{\hat{d} - \hat{c}}$$

- Results can be transferred to dose-reponse functions of the form:

$$f(x; c, d, EC50, \ldots) = c + (d - c)\, h\left(\frac{x}{EC50}, \ldots\right)$$

References

- Ritz, C. (2010). Toward a unified approach to dose–response modeling in ecotoxicology. *Environmental Toxicology and Chemistry* **29**, 220–229.
- Ritz, C., Streibig, J.C. (2005) Bioassay analysis using R. *Journal of Statistical Software* **12**, 1–22.
- Weimer M, Jiang X, Ponta O, Stanzel S, Freyberger A., Kopp-Schneider A. (2012) The impact of data transformations on concentration-response modeling. *Toxicology Letters* **213**, 292– 298.