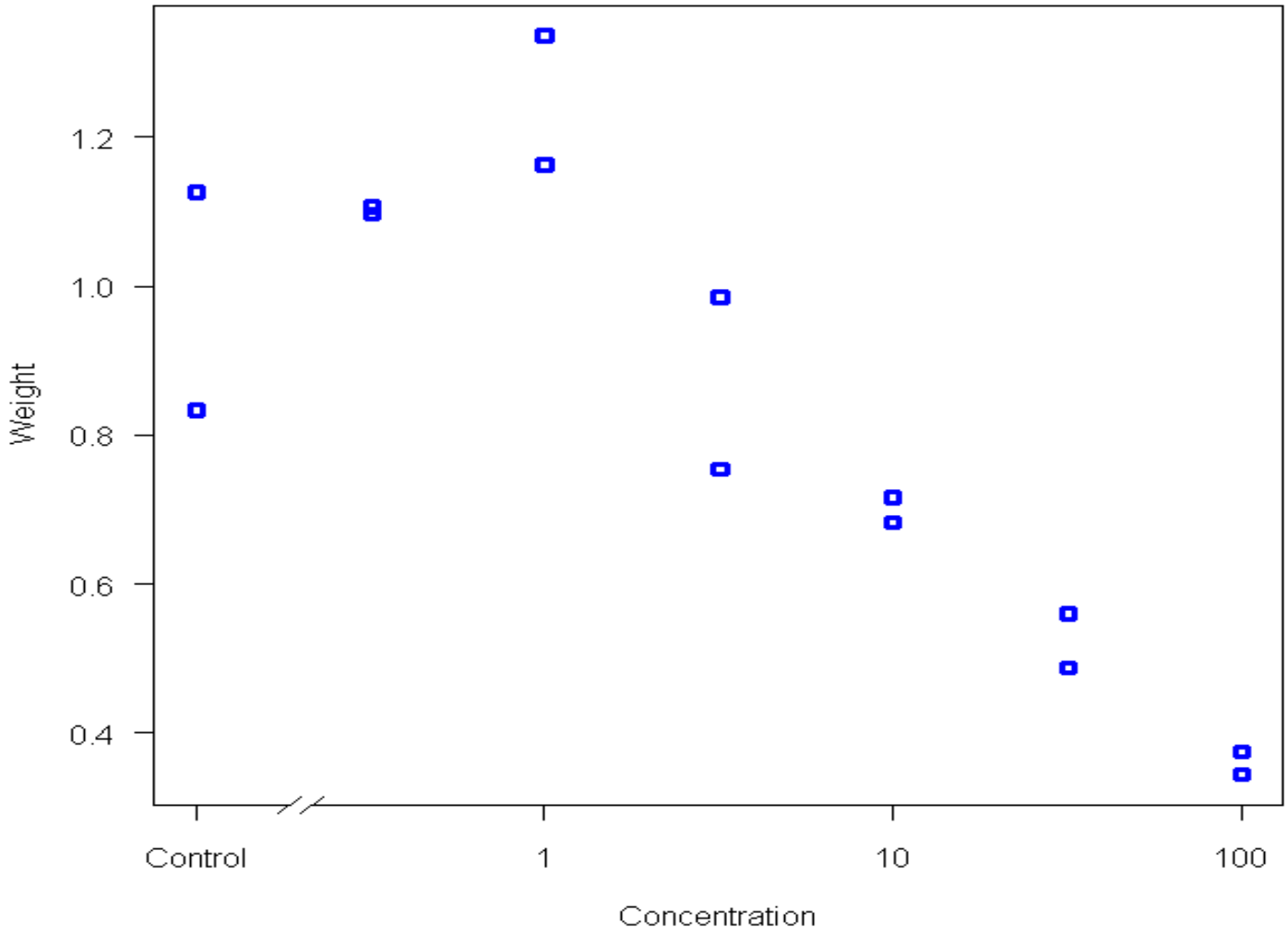
Use of model averaging in dose-response analysis

Christian Ritz

Statistics Group
Department Basic Sciences and Environment
Faculty of Life Sciences
University of Copenhagen

NCS 2010, Lyon, September 28 2010

Data example – which parametric model to use?



Why model averaging?

About a single parametric model fit:

+ computationally fast and reliable nowadays
– strong assumptions about the expected dose-response pattern
(we need to decide in advance on the appropriate model!)

Model averaging of several parametric models:

- our ignorance expressed by providing a pool of candidate models
- model selection *uncertainty* not ignored!
- computationally feasible in an automatic fashion
- slightly larger uncertainty on estimates – so it should be!

Weighting of models

Use a *relative* measure of the merits of each individual model!

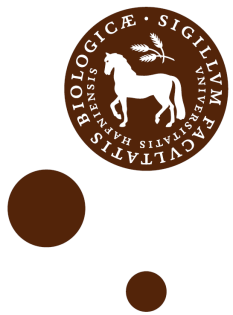
Some available measures (so-called information criteria):

- Akaike's information criterion (AIC)
  - *smallest distance between candidate models and the true but unknown model*
- Bayesian information criterion (BIC)
  - *largest posterior probability among candidate models*

BIC penalizes complex models more severely than AIC
– not always useful for small data sets!

Weights calculated using the formula (for M models):

$$w_i = \frac{\exp(-IC_i/2)}{\sum_{j=1}^{M} \exp(-IC_j/2)}$$



Model-averaged estimates

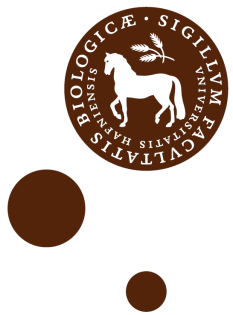
Estimated effect concentration defined as follows (*conditional on the weights*):

$$\widehat{EC}_{MA}(p) = \sum_{i=1}^{M} w_i \widehat{EC}_i(p)$$

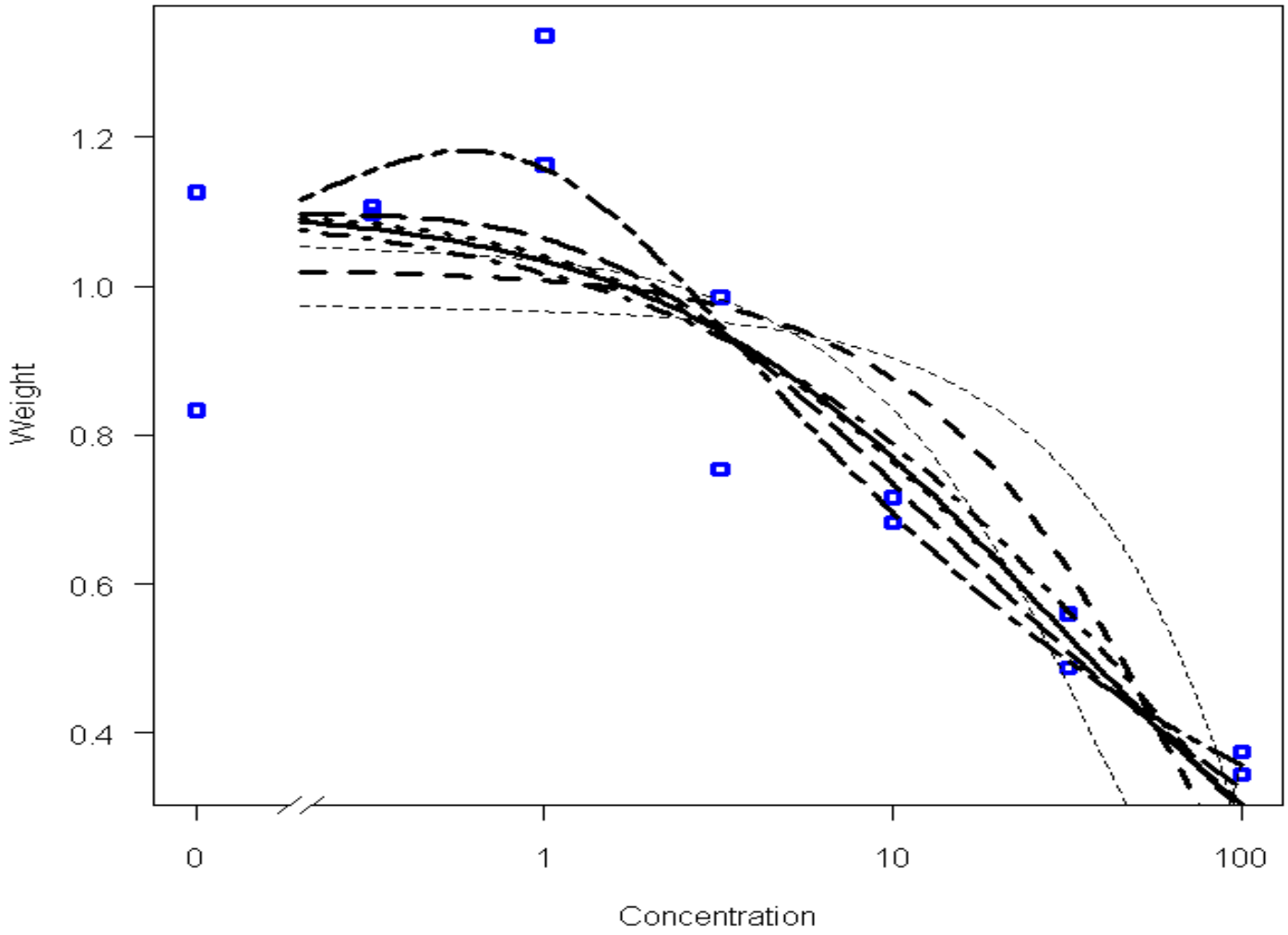
($\widehat{EC}_i(p)$ estimated EC value for model i)

Confidence limits of EC values:

- Average individual confidence limits – completely *ad hoc*! (Kang *et al.*, 2000; Wheeler & Bailer, 2009)
- Assume perfect correlation between individual estimates – not all information used, conservative (Buckland *et al.*, 1997)



Data example revisited I



Data example revisited II

Output from **R** *(best model fit in the top)*:

```
        log Lik        AIC  Res. var.      Weights
BC.4  13.165753 -16.331505 0.01249750 0.8211333256  - Brain-Cousens
W2.3   9.843867 -11.687734 0.01826111 0.0805436982
LN.3   8.923033  -9.846066 0.02082848 0.0320714284
Cubic  9.549023  -9.098045 0.02095138 0.0220641779  - cubic
LL.3   8.533544  -9.067088 0.02202025 0.0217252829  - log-logistic
Quad   7.826190  -7.652380 0.02436172 0.0107094094  - quadratic
W1.3   7.790219  -7.580439 0.02448723 0.0103310312
L.3    5.480355  -2.960710 0.03406030 0.0010256100
Lin    3.528818  -1.057636 0.04126076 0.0003960364  - linear
```

Model-average EC50 estimate:

```
     Estimate Std. Error    Lower    Upper
1:50 33.89621   14.97240 4.550839 63.24158
```

Reference EC50 estimate based on standard log-logistic:

```
     Estimate Std. Error   Lower  Upper
1:50  28.6064    10.8569  4.7106 52.502
```

Concluding remarks

Key points:

- Powerful approach for unsupervised dose-response analysis of high-throughput data
- Inclusion of polynomial regression to capture unusual patterns
- Implemented in **R**

Challenges:

- Establishing a comprehensive and flexible pool of models *(Avoid nested models!)*
- Improved methods for evaluating the uncertainty of model-averaged estimates
- More robust automatic starting value procedures desirable

References

- Buckland, S. T., Burnham, K. P. & Augustin, N. H. (1997). Model selection: An integral part of inference. *Biometrics* **53**, 603–618
- Kang, S.- H., Kodell, R. L. & Chen, J. J. (2000). Incorporating Model Uncertainties along with Data Uncertainties in Microbial Risk Assessment. *Regulatory Toxicology and Pharmacology* **32**, 68–72
- Wheeler, M. W. & Bailer, A. J. (2009). Comparing model averaging with other model selection strategies for benchmark dose estimation. *Environmental and Ecological Statistics* **16**, 37–51