Comparison of initial value routines for dose-response models

Anke Schulz

September 28^{th}, 2010
NCS 2010 in Lyon

Bayer HealthCare
Bayer Schering Pharma

Motivation

- Dose-response models are routinely used in different areas and often run in an automated manner.
- Dose-response relations are often described by sigmoid models
- The non-linearity of the sigmoid model requires an iterative parameter estimation procedure.
- The convergence behavior of the iteration procedure depends on the choice of starting values.

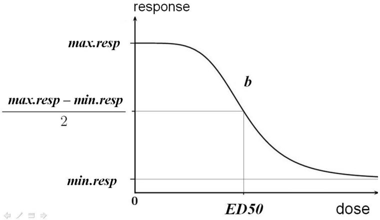
Dose-response model

$$f(x) = y = min.resp + (max.resp - min.resp) \cdot g\left(x, (b, ED_{50})\right)$$

- for the basic types of sigmoid models (e.g. 4 parameter logistic, Log-normal, Weibull)

$$g\left(x, (b, ED_{50})\right) = \tilde{g}(b \cdot \log(x) - b \cdot \log(ED_{50})).$$

y	:	measured response
x	:	corresponding dose
$min.resp$	:	lower horizontal asymptote
$max.resp$	:	upper horizontal asymptote
$g()$	:	sigmoid curve
ED_{50}	:	dose of the middle response
b	:	slope related



e.g., 4 parameter logistic model: $y = min.resp + \dfrac{(max.resp - min.resp)}{1 + \left(\frac{X}{ED_{50}}\right)^{b}}$

Starting value routines

- Regression approach (Ritz and Streibig 2005)
- Normolle (1993)
- New idea

- For all routines, the starting values for the asymptotes are:

$$
\begin{aligned}
max.resp_{st} &= \max_i(y_i) + 0.01 \cdot \left(\max_i(y_i) - \min_i(y_i)\right), \\
min.resp_{st} &= \min_i(y_i) - 0.01 \cdot \left(\max_i(y_i) - \min_i(y_i)\right), \\
& i = 1, \ldots, n
\end{aligned}
$$

Starting value routines - Regression approach

- Transform the model to get an equation with a linear term in $\log(x)$ on one side.

$$y = min.resp_{st} + (max.resp - min.resp) \cdot \tilde{g}(b \cdot \log(x) - b \cdot \log(ED_{50})).$$

$$\tilde{g}^{-1}\left(\frac{y - min.resp}{max.resp - min.resp}\right) = b \cdot \log(x) - b \cdot \log(ED_{50}).$$

- Insert starting values for *min.resp* and *max.resp*.
- Parameter estimation of the regression model.

Starting value routines - Normolle

- Starting value for ED_{50}: Middle of dose groups.

$$ED_{50\ st} = \min_i(x_i) + 0.5(\max_i(x_i) - \min_i(x_i)).$$

- Solve the model equation by b and calculate b_i for each observation.

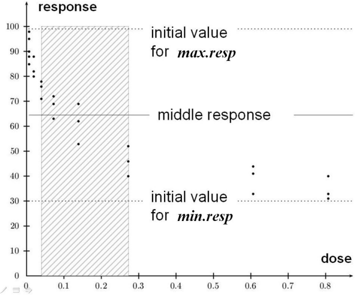
$$\begin{aligned} b_i &= \frac{\tilde{g}^{-1}\left(\frac{y - min.resp_{st}}{max.resp_{st} - min.resp_{st}}\right)}{\log(x_i) - \log(ED_{50\ st})}; \\ b_{st} &= \text{median}(b_1, \ldots, b_n); \quad i = 1, \ldots, n \quad \text{and} \end{aligned}$$

$$\begin{aligned} ED_{50\ i} &= exp\left(\frac{b\log(x) - \tilde{g}^{-1}\left(\frac{y - min.resp_{st}}{max.resp_{st} - min.resp_{st}}\right)}{b_{st}}\right); \\ ED_{50\ st} &= \text{median}(ED_{50\ 1}, \ldots, ED_{50\ n}); \quad i = 1, \ldots, n \end{aligned}$$

Starting value routines - New approach
Starting value for ED_{50}

Dose interval für the starting value of ED_{50}

- Middle response:
  $mr = \frac{max.resp - min.resp}{2}$
- Dose interval=[biggest dose with all observations above mr ; smallest dose with all observations below mr].
- $ED_{50\ st}$: weighted mean of the observations of the interval. Weight depends on the distance to the middle response.



Starting value routines - New approach
Starting value for b

- the first derivative at $x = ED_{50}$ is used for the starting value for b.

$$y' = f'(x) = (max.resp - min.resp) \cdot g'\left(x, (b, ED_{50})\right)$$

- Solve $f'(x)$ by b.
- $x = ED_{50}$; $f'(ED_{50})$ = slope of the lin. reg. of the dose interval

e.g., 4 parameter logistic model:

$$f'(ED_{50}) = \frac{-(max.resp - min.resp) \cdot b}{4 \cdot ED_{50}}$$

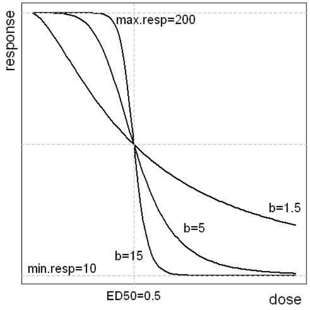
$$b = \frac{-4 \cdot ED_{50\ st} \cdot f'(ED_{50\ st})}{max.resp_{st} - min.resp_{st}}.$$

Simulation

- Model: four-parameter logistic regression model by varying the parameters ED_{50} and b and the variance of the simulated data and $min.resp = 10$ and $max.resp = 200$.

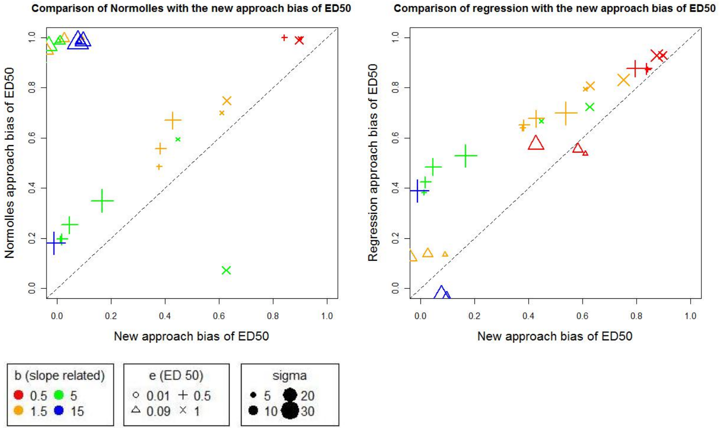
$$
\begin{aligned}
x &= \{0.0, 0.025, 0.05, 0.1, 0.2, 0.4, 0.8\} \\
ED_{50} &\in \{0.01, 0.09, 0.5, 1\} \\
b &\in \{0.5, 1.5, 5, 15\} \\
\sigma &\in \{5, 10, 20, 30\}
\end{aligned}
$$

- $N = 10000$ runs



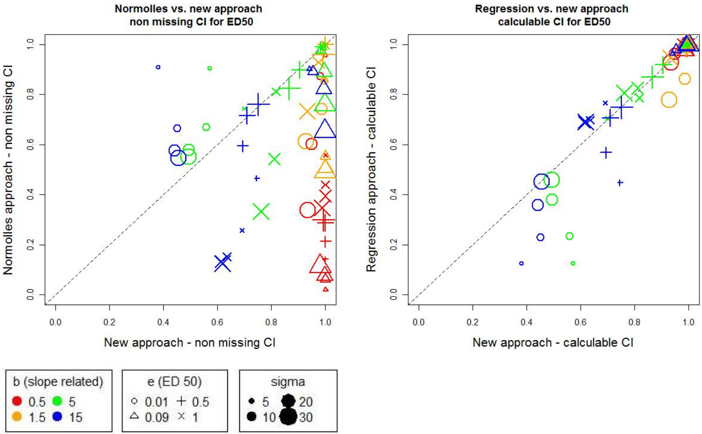
Results

- Bias of ED_{50}.
- Comparison of Normolle vs. New and regression vs. New.



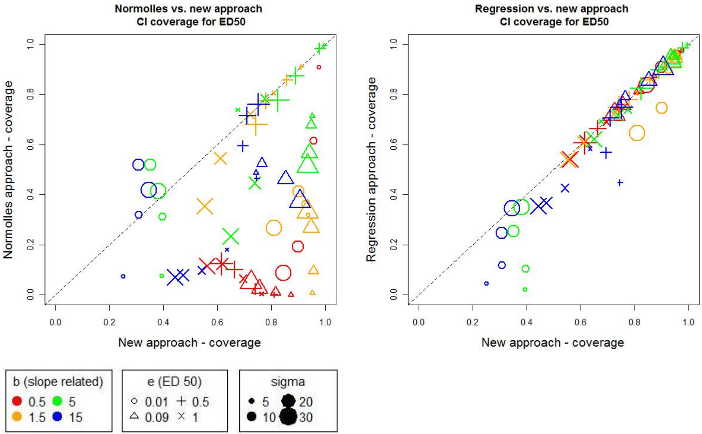
Results

- Non missing confidence intervals for ED_{50}.
- Comparison of Normolle vs. New and regression vs. New.



Results

- Coverage of the confidence interval of ED_{50}.
- Comparison of Normolle vs. New and regression vs. New.



Conclusion

In general

- All starting value routines perform best for the ideal design (i.e. curve with moderate slope and ED_{50} in the middle of dose range).
- Variation in the data has very low impact on iteration performance compared to the design.

Routine comparison

- The New approach is superior to Normolle in all cases except for the ideal design.
- The New approach performs better or at least not worse than regression approach in all designs.
- The New approach is especially preferable if reduction of bias is most relevant or if ED_{50} is suspected at either end of the dose range.

References

- Normolle, D.P.(1993): An Algorithm for Robust non-linear Analysis of Radioimmunoassays and other Bioassays.
  Statistics in Medicine, 12: 2025-2042.
- Ritz, C. and Streibig, J.C. (2005): Bioassay Analysis using R.
  Journal of Statistical Software, 12, Issue 5.
- Schulz A. (2006): Das vierparametrische logistische Modell und seine Anwendung bei einem ökotoxikologischen Problem.
  Master Thesis, Humboldt-Universität zu Berlin, Germany.