Challenges of benchmark dose analyses for risk assessors in the wildlife risk assessment

Christian Dietrich[#], Magnus Wang[#]

[#]WSC Scientific GmbH, Heidelberg, Germany - [E-mail contact: magnus.wang@wsc-regexperts.com]

Abstract

Benchmark dose (BMD) analyses are now implemented as a routine method in the risk assessment scheme of birds of mammals (EFSA, 2023). Several software tools are proposed for BMD calculations and a guidance has been developed on how to conduct calculations (EFSA, 2022). However, while the guidance and software tools seem to make BMD calculations an easy routine task, there are a number of pitfalls, which should be considered. In this poster, we discuss some of the most urgent issues, including new issues introduced by the recent development of BMD methodology.

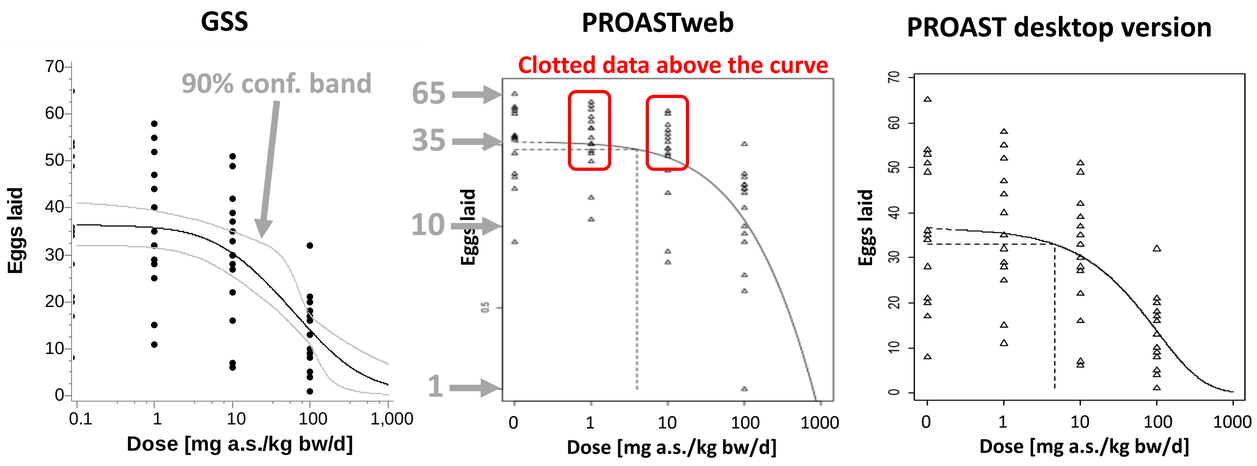
Introduction

For benchmark dose (BMD) modelling, being a parametric statistical method, certain assumptions regarding data distribution or homogeneity of variances are made, which are the basis for obtaining correct results. Often, the fulfilment of these prerequisites is insufficiently checked or not checked at all. For example, in the case of biological data used for benchmark dose analyses, it is often not clear how they are actually distributed. Incorrect assumptions about the data distribution or even the selection of models that do not make biological sense can lead to the results of the analysis being significantly distorted. In the following, we show examples.

Methods

To evaluate results from different benchmark dose software realistic continuous individual data were generated as they are frequently reported in typical bird and mammal reproduction studies. BMD analyses were carried out using the commercial software GSS (WSC Scientific GmbH), PROASTweb, the PROAST desktop version, and the Bayesian BMD tool (Laplace Approximation LA and Bridge Sampling BS).

Results for avian reproduction endpoint ‘eggs laid’



It can be seen that data are approximately normally distributed. Hence, no transformation is needed before fitting.

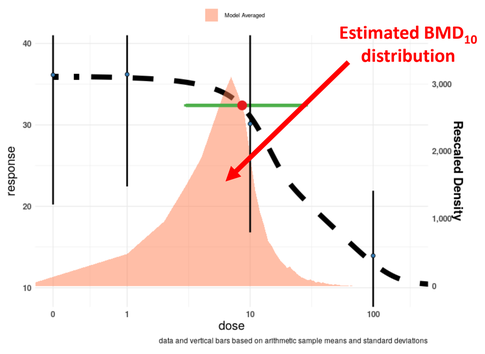
Figure 1. ’Average model’ calculated with GSS.

Data were automatically log-transformed (this can not be turned off in PROASTweb). The data are no longer normally distributed and distort the model fit.

Figure 2. Model fit from PROASTweb.

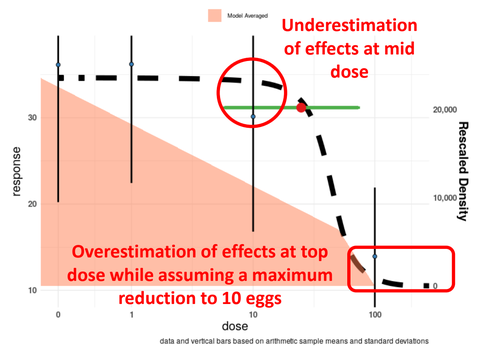
Data transformation can be turned off in the desktop version of PROAST. However, no model averaging is available when transformation is disabled.

Figure 3. Model fit from PROAST desktop version.



Although the ‘average model’ appears to adequately describe the data, the BMD_{10} exceeds the BMDs of GSS and PROAST by more than 50% and the BMDL exceeds the BMDLs of GSS and PROAST by a factor of five (see Table 1).

Figure 4. ‘Average model’ calculated with the Bayesian BMD tool using Laplace approximation.



Using bridge sampling resulted in very bad model fits. Although the mid dose already shows more than 10 % effect on average, the BMD_{10} was estimated significantly higher than this dose and much higher than BMDs from other software/methods (see Table 1).

Figure 5. ‘Average model’ calculated with the Bayesian BMD tool using bridge sampling.

When comparing the results from the different approaches and tools (Table 1), it can be seen that results vary. While the results from PROAST and GSS were comparable, the BMD_{10} calculated by the Bayesian BMD tool were much larger. In particular, the approach using bridge sampling resulted in 50 times higher BMD_{10} values.

References:

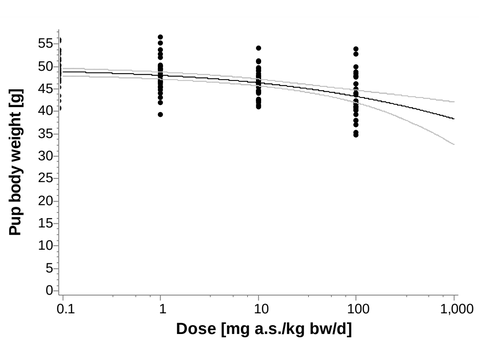
EFSA 2022. Guidance on the use of the benchmark dose approach in risk assessment. EFSA J. 20: 7584. https://doi.org/10.2903/j.efsa. 2022.7584

EFSA 2023. Guidance on the risk assessment for Birds and Mammals. EFSA J. 21: 7790. https://doi.org/10.2903/j.efsa.2023.7790

Software	Figure	Data log-transformed	Model averaging applicable	BMD_{10}	$BMDL_{10}$ / $BMDU_{10}$
GSS	Figure 1	Optional (not used)	Yes	5.50	0.52 / 43.00
PROASTweb	Figure 2	Always (automatically)	Yes	No estimate provided	0.21 / 47.40
PROAST desktop*	Figure 3	Optional (not used)	Only for transformed data	4.63	0.35 / 57.20
Bayesian BMD tool (LA)[#]	Figure 4	Optional (both used)	Yes	8.60 8.61	3.00 / 28.52 2.94 / 28.40
Bayesian BMD tool (BS)[#]	Figure 5	Optional (both used)	Yes	24.83 247.85	5.83 / 73.07 7.48 / 868.34

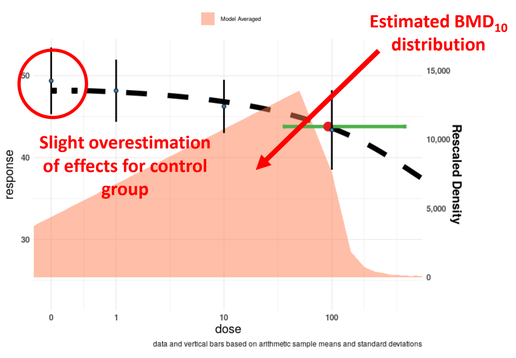
Table 1. Comparison of different BMD software evaluating the avian endpoint ‘eggs laid’. (*) Note that for the PROAST desktop version the BMD and its confidence interval were not based on model averaging but refer to the best fitting model (lowest AIC among all fitted models). (#) Evidence against log-normality of the data triggered an additional analysis with adapted prior weights giving a second BMD estimate (top value: BMD from default calculations, bottom value: BMD from calculations with adapted weights).

Results for mammalian reproduction endpoint ‘pup body weight’



The relatively low variance in the data and the large sample size (n=30) results in narrow confidence bands and reliable estimates for the BMD_{10} and its associated confidence interval.

Figure 6. ’Average model’ calculated with GSS.



Since the ‘average model’ already predicts slight effects in the control group (red circle), 10% effect is estimated at higher dose. The BMD_{10} exceeds the BMDs of GSS and PROAST by more than 80% (see Table 2).

Figure 7. ’Average model’ calculated with the Bayesian BMD tool using bridge sampling.

Software	Figure	Data log-transformed	Model averaging applicable	BMD_{10}	$BMDL_{10}$ / $BMDU_{10}$
GSS	Figure 6	Optional (not used)	Yes	48.67	17.87 / 163.92
PROASTweb	Not shown	Always (automatically)	Yes	No estimate provided	14.90 / 113.00
PROAST desktop*	Not shown	Optional (not used)	Only for transformed data	50.08	16.80 / 132.00
Bayesian BMD tool (LA)	Not shown	Optional (both used)	Yes	62.83	26.08 / 138.91
Bayesian BMD tool (BS)	Figure 7	Optional (both used)	Yes	92.56	36.32 / 473.96

Table 2. Comparison of different BMD software evaluating the mammalian endpoint ‘pup body weight’. (*) Note that for the PROAST desktop version the BMD and its confidence interval were not based on model averaging but refer to the best fitting model (lowest AIC among all fitted models).

Issues with zeros

PROASTweb and the Bayesian BMD modelling tool cannot handle zeros. Therefore, zeros have to be removed or changed to a small value (e.g. 0.01). The choice of the small value has great impact on model results and whether the zeros are identified as outliers (the smaller the value to be log-transformed, the further away it is from the actual data set and the higher the probability that the value will be identified as an outlier).

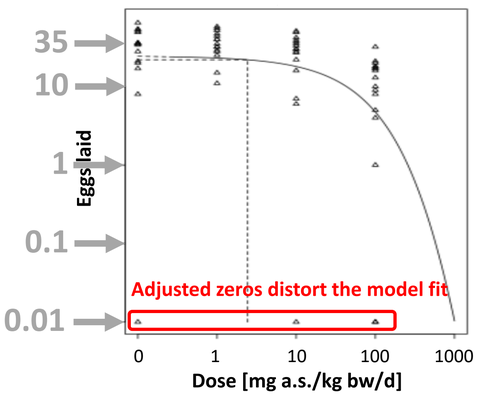


Figure 8. Model fit from PROASTweb illustrating issues arising from log-transforming adjusted zeros.

Conclusion

Overall, the BMD approach is a useful tool for estimating the dose-response relationship and identifying a reference point for the risk assessment, but it is not without conceptual issues and potential pitfalls. Careful consideration of the assumptions is necessary to obtain reliable results.

Since different software sometimes produces very different results for the same data, BMD analysis should not be limited to a single software in justified cases (e.g. when model fits do not visually fit the data well). The use of additional software and a worst-case approach to the results could be considered.