

Manual for BMD Modeling

Contents

Introduction	1
Installation	1
Getting started	2
About and Report new issue	2
Methodology	3
Fitting the BMD models using Proast	3
Model averaging	3
Data	3
Fit Models	6
Analysis for quantal response	6
Single model for quantal response	8
Analysis for quantal response with covariate	8
Analysis for continuous summary response	9

Introduction

This online application implements statistical methods for Benchmark dose modeling using Proast (version 61.3) and averaging results from multiple fitted benchmark dose models. It allows to estimate the dose that corresponds with the benchmark response of interest. The estimated benchmark dose (BMD) is reported along with its lower and upper bound. When fitting a list of models, the weighted average of the model-specific BMD estimates can be obtained for quantal response type data. The lower and upper bound for the BMD are then estimated using parametric bootstrap sampling.

Installation

The application is available at <https://efsa.openanalytics.eu/app/bmd>, provided you have an account.

Alternatively, the application can be run locally by installing the package on your computer.

Prior to the installation, please make sure that the proxy is configured properly to access the internet from EFSA premises:

```
library(httr)
set_config(use_proxy(url="tmgproxy", port=8080, username="user",
                    password="password"))
```

To run the application, you will first need to install the most recent version of the package “bmdModeling” in R. Then load the package and start the application with the runBmd() function.

`library(bmdModeling)`

`runBmd()`

Getting started

This is the start screen of the online application:

Figure 1: Start screen for bmd.

Using the tab pages, one can switch between:

- **Data:** Specify the data that should be used for analysis.
- **Fit Models:** Fit benchmark dose models using Proast and perform model averaging if available.

After the analysis is performed, a Word document with summarized results can be obtained by pushing the button **Download report** at the top of the page.

About and Report new issue

The user is given a summary of the application’s functionalities when clicking on the link “About” at the right top. New issues can be reported via “Report new issue”. You are then redirected to the EFSA Model Manager website and invited to describe the issue.

Methodology

Fitting the BMD models using Proast

Given the dose-response data, the shape of the overall dose-response relationship is estimated for a particular endpoint. The BMD (benchmark dose) is a dose level, estimated from the fitted dose-response curve, associated with a specified change in response, the Benchmark Response (BMR). The BMDL and the BMDU are respectively the lower and upper confidence bound of the BMD. The benchmark dose models are fitted using the R-package `proast61.3`.

Using this application one can either fit a single model or a set of models to the data. In the latter case, we focus on the accepted models, i.e. models for which the $AIC \leq AIC_{Min} + 2$ as shown in the flowchart (Figure 2). The lowest BMDL value of all accepted models is normally used as the Reference Point (RP), unless the model averaged BMDL is available, see Model averaging.

Model averaging

Model averaging is currently only implemented for non-continuous response type data.

The model averaging principles are explained in detail in Wheeler and Bailer (2007). Consider a set of models where each model estimates the probability of the response event to happen. Then there are two approaches to estimate the model averaged BMD:

1. Calculate a weighted average of each model's estimated BMD or
2. Construct a weighted average model, which weights the estimated response probabilities.

In both cases weighing is based on the model's AIC values where better models get larger weight. For the first naive approach, an estimate for the model averaged BMD is directly obtained. For the second approach, the inverse of the model averaged response curve need to be estimated in order to obtain an estimate of the model averaged BMD given the BMR.

In both cases, the Reference Point is the model averaged BMDL, the lower bound of the model averaged BMD which is estimated using parametric bootstrap.

Data

This tab page guides the user in loading and specifying the data for analysis. The data file can be uploaded using the **Browse** button.

- The **Data format** link explains the structure of the two data types that can be uploaded: "Proast data" or "Raw data".
- Select the **Data separator** that separates data values. This can be either 'Comma'/'Semicolon' mostly for .csv files, or 'Tab'/'Space' mostly for .txt files.
- Select the **Decimal separator** that defines the decimal delimiter. This can be either 'Point' or 'Comma'.
- The user can select a variable name for **Subset of the data according to**, which will then keep only those records that have the **keep value**. If "none" is selected no subset will be taken (default).

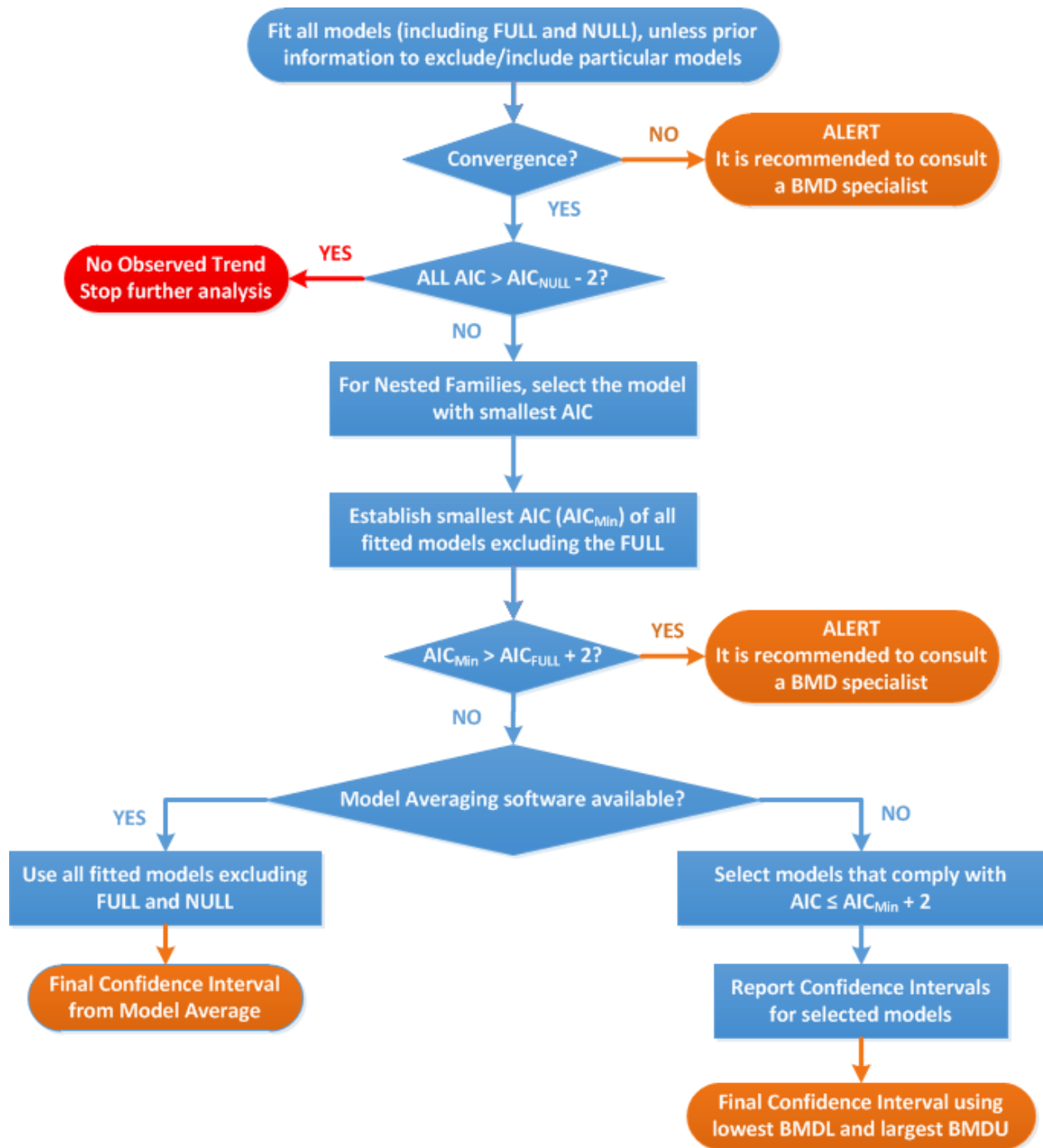


Figure 2: Flowchart for selection of BMDL

- In the field **Which response(s) do you want to consider?** one can select the variable names of the responses to be used in further analysis. The **List of non-responses** will then automatically display the other variables of the loaded data.
- The **Type of response** can be either ‘continuous’, ‘quantal’, ‘binary’ or ‘ordinal’. Depending on this choice, the user may be asked to further specify whether the responses are reported at individual or summary level and whether there is a litter effect or not. These specifications determine e.g. which models will be available in the “Fit Models” tab and which model parameters need to be specified.

In Figure 3 quantal response data are loaded (proast data format). Continuous summary data in raw data format are loaded in Figure ???. Both files, respectively “methyleug.txt” and “das10.csv” are available in “/bmdModeling/inst/extdata”.

The screenshot shows the 'Data' tab with 'Fit Models' selected. The file 'methyleug.txt' is uploaded. The 'Data format' section is set to 'Option 1: Proast data'. The 'Data separator' is 'Tab (.txt file)'. The 'Decimal separator' is 'Point'. The 'Subset of the data according to:' is 'none'. The 'Which response(s) do you want to consider?' field contains 'forest.kk liver.kk'. The 'List of non-responses:' is 'dose.kg.bw, sample.size, sex'. The 'Type of response' is 'quantal'. The 'Litter effect' checkbox is unchecked.

dose.kg.bw	forest.kk	liver.kk	sample.size	sex
0	0	0	50	1
3	8	4	50	1
10	40	10	50	1
30	50	15	50	1
0	1	0	50	2
3	6	2	50	2
10	30	15	50	2
30	49	38	50	2

Figure 3: Load quantal data in proast data format.

The screenshot shows the 'Data' tab with 'Fit Models' selected. The file 'das10.csv' is uploaded. The 'Data format' section is set to 'Option 2: Raw data'. The 'Data separator' is 'Comma (.csv file)'. The 'Decimal separator' is 'Point'. The 'Subset of the data according to:' is 'none'. The 'Which response(s) do you want to consider?' field contains 'mean'. The 'List of non-responses:' is 'ppm, dose, sex, sd, number, X2'. The 'Type of response' is 'quantal'. The 'Litter effect' checkbox is unchecked.

ppm	dose	sex	mean	sd	number	X2
0	0	1	704	124.7	33	21.7074836976755
10	0.5	1	739	140.5	35	23.7488345576713
70	3.5	1	742	97.7	40	15.447263698225
500	25	1	646	119.4	41	18.6471471694996
1000	50	1	572	97	49	13.8571428571429
0	0	2	496	105.7	37	17.3769729581763
10	0.5	2	477	132.6	33	23.0626971797255
70	3.5	2	480	106.8	32	18.8797510576808
500	25	2	402	106.8	27	20.5536695831507
1000	50	2	361	81.1	20	18.1345112975233

Figure 4: Load continuous summary data in raw data format.

At the right hand side it can then be checked whether the data were uploaded correctly. If not, you may want to adapt the data before uploading or you may have to change the data/decimal separator used for loaded data in the left panel. The shown table has several features:

- The shown number of entries can be changed.
- Specific sampling items can be searched for in the field at the right top.
- Data can be ordered according to a given characteristic by clicking on its name.

Fit Models

The input fields in this tab page will differ depending on the chosen type of response. In this section we will discuss some examples for benchmark dose modeling.

Analysis for quantal response

In Figure 5, the default settings for the loaded quantal data are shown. At the left hand side, the user is asked to select the independent variable, response variable and sample size for each of the selected response variables. When selecting multiple response variables of the same type, results are printed per response.

It is recommended to start by fitting a **Set of models**. Afterwards, when fitting problems are experienced, a single model can be selected to fit it with modified model parameters constraints. For non-continuous response type models, it is recommended to **Perform model averaging** (selected by default). By clicking on the advanced settings link, the user can change the default settings for model averaging:

- Include results for the naive approach: the model averaged BMD is estimated as a weighted average of all model's estimated BMD values (see section 'Methodology: Model averaging').
- By default all converged models are considered for model averaging, but for each response the user can change the set of models to be considered.

At the right hand side, the Additional settings tab identifies:

- The **Benchmark criterion**: For quantal response data three choices are available:
 - ED50: The BMD (benchmark dose) is defined as the dose that corresponds with an estimated risk of 50% (ED50).
 - Additional risk: The BMD is defined as the dose that corresponds with an additional risk of e.g. 5% compared with the background risk.
 - Extra risk: The BMD is defined as the dose that corresponds with an extra risk of e.g. 5% compared with the background risk (Default option).
For the latter two the value of 0.05 (for continuous response) or 0.1 (for non-continuous response) is specified in the input field **Value for the BMR**.
- A value of 0.9 for the **Confidence level for the BMD confidence intervals** specifies that the reported BMDL and BMDU are respectively the estimated lower and upper bound of the 90% confidence interval for the BMD.
- When confidence intervals are constructed using bootstrap sampling, an integer value for the **The number of bootstrap runs for calculating BMD confidence intervals** is required (default 200).

For the Model parameters tab, we use the default settings. When all settings are set, push the button **Fit Model(s)**.

The screenshot shows the 'Fit Models' tab in the Proast software. It is divided into two main sections: 'Define variables' and 'Additional settings / Model parameters'.
 In the 'Define variables' section:
 - 'Independent variable (e.g. dose)' is set to 'dose.kg.bw'.
 - 'Response variable(s)' is set to 'forest.kk'.
 - 'Sample size for response variable:' is set to 'sample.size'.
 - 'Models to fit' is set to 'Set of models'.
 - There is a checked box for 'Perform model averaging' and a link for 'Advanced settings'.
 In the 'Additional settings / Model parameters' section:
 - 'Benchmark criterion' is set to 'Extra risk, i.e. (P(BMD) - P(0)) / (1 - P(0))'.
 - 'Value for the BMR, in terms of extra risk' is set to 0.1.
 - 'Confidence level for the BMD confidence intervals' is set to 0.9.
 - 'The number of bootstrap runs for calculating BMD confidence intervals' is set to 200.
 At the bottom right, there is a small text: 'Output generated with Proast, version 61.3'.

Figure 5: Quantal response data: default input values.

The output of the fitted models is shown in Figure 6. The list of fitted models differs for continuous and non-continuous response type data. In the table the estimated BMD along with its 90% confidence limits BMDL and BMDU are reported. It is also indicated whether the model converged and whether the model was accepted based on its AIC value and the flowchart in Figure 2. The table can be saved as “.csv” file by pushing the button “Download”.

Response: forest.kk

Fitted Models

Model	Included covariate(s)	Number of parameters	Log-likelihood	AIC	BMD	BMDL	BMDU	Converged	Accepted AIC
Null		1	-275.98	553.96	NA	NA	NA	yes	
Full		4	-112.78	233.56	NA	NA	NA	yes	
Logistic		2	-117.75	239.50	3.35	2.84	3.94	yes	no
Probit		2	-122.32	248.64	3.03	0.00	3.54	yes	no
Log-logistic		3	-113.59	233.18	2.80	2.21	3.51	yes	yes
Log-probit		3	-113.01	232.02	2.76	2.22	3.34	yes	yes
Weibull		3	-113.81	233.62	2.12	1.54	2.77	yes	yes
Gamma		3	-113.05	232.10	2.45	1.78	3.16	yes	yes
Two-stage		3	-115.06	236.12	1.82	1.22	2.78	yes	no

Download

Figure 6: Quantal response data: fitted models.

For non-continuous response type models, results for model averaging are available, as shown in Figure 7. The estimated model weights and the model averaged BMD, BMDL and BMDU are reported. The averaged response model is plotted in black, based on the original data and from which the model-averaged BMD is estimated. The gray lines indicate the averaged response models based on the bootstrap data, from which the BMDL and BMDU are estimated. This plot can be saved as “.png” file by pushing the button “Download”.

All reported results can directly be copied into a Word document by pushing the button **Download**

report at the top of this page.

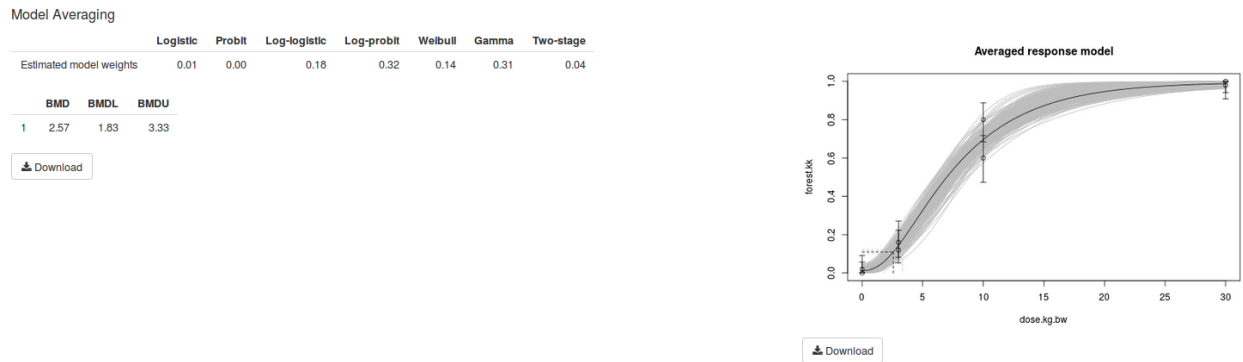


Figure 7: Quantal response data: model averaging.

Single model for quantal response

When for a specific model there was no convergence or acceptance, we may prefer to refit it with adapted constraints for the model parameters. We fit the two-stage model with adapted values for the model parameters' lower and upper bound and start values (see Figure 8). These options can be found in the advanced settings of the Model parameters tab. The results remain the same, but now also a plot of the two-stage model is displayed, where we can see that the model does not fit well for some dose values.

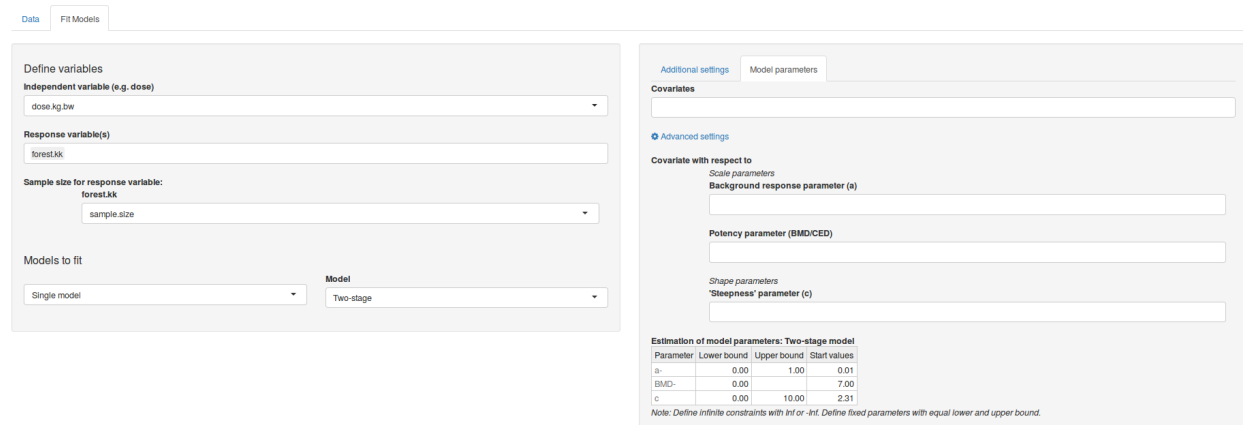


Figure 8: Quantal response data: single model.

Analysis for quantal response with covariate

We fit again a set of models, but select in the Model parameters tab 'sex' for **Covariates**. This will by default fit models including a separate parameter value per level of sex for:

- the background response parameter (a), or
- the potency parameter (BMD/CED), or

- the background response parameter (a) and the potency parameter (BMD/CED), or
- none of the parameters.

Among these four models, the one with the lowest AIC value is reported. The user can ask to **Show results for all covariate combinations**. It is not recommended to change the default settings for which parameters covariates are selected or not, but if desired the user can (de)select covariates for individual parameters in the Advanced settings (Figure 9).

The screenshot shows a web-based interface for fitting models. On the left, the 'Define variables' panel includes:

- Independent variable (e.g. dose):** A dropdown menu with 'dose.kg.bw' selected.
- Response variable(s):** A text input field containing 'forest.kk'.
- Sample size for response variable:** A dropdown menu with 'sample.size' selected.
- Models to fit:** A dropdown menu with 'Set of models' selected.

 Below this panel are two checkboxes: 'Perform model averaging' (checked) and 'Advanced settings' (with a gear icon).

On the right, the 'Additional settings / Model parameters' panel includes:

- Covariates:** A text input field containing 'sex'.
- Show results for all covariate combinations
- Advanced settings:** A blue gear icon with the text 'By default no covariate is selected for 'Steepness' parameter (c) and the input field is blank.'
- Covariate with respect to:** A dropdown menu with 'sex' selected.
- Scale parameters:**
 - Background response parameter (a):** A text input field containing 'sex'.
 - Potency parameter (BMD/CED):** A text input field containing 'sex'.
- Shape parameters:**
 - 'Steepness' parameter (c):** An empty text input field.

 At the bottom of this panel, a note states: 'Default settings for estimation of model parameters can only be changed when a single model is being fit'.

Figure 9: Quantal response data with covariate: input values.

For each of the models, the BMD, BMDL and BMDU is reported per level of the selected covariate for BMD, here sex (Figure 10). We observe that for the probit model currently no covariates can be included, for the logistic model the best one is including a covariate for parameter a, while for the others the best one is including a covariate for parameter b only. In general the estimated BMD is lower for level 1 than level 2. Notice that the number of model parameters has increased compared to the fitted models without covariates (Figure 6). As before, results for model averaging are reported, the model averaged BMDL is then used as reference point. It is clear that a separate curve per level of sex fits well to the data.

When selecting the option to show results for all covariate combinations, the results for all fitted models per model family are reported (Figure 11).

Analysis for continuous summary response

Consider the continuous summary data loaded in Figure 4. We select 'ppm' as independent variable and 'mean' as response variable. The **value for CES** is set to 0.05 stating that the BMD will correspond with an estimated difference in response of 5% compared with the background response (when ppm is 0). For the Model parameters tab, we use the default settings: No covariates are selected.

Results of the fitted models are shown in Figure 13. Note that the set of fitted models is different from that for quantal responses e.g. in Figure 6. None of the models is accepted following the rules in the flowchart (Figure 2). Therefore it is recommended to consult a BMD specialist. As there is

Fit Model(s)

Response: forest.kk

Fitted Models

Error(s) in calculation:

• Probit model: Factor for parameter 'a' or 'b' not implemented for probit model

Model	Included covariate(s)	Number of parameters	Log-likelihood	AIC	BMD - sex_1	BMDL - sex_1	BMDU - sex_1	BMD - sex_2	BMDL - sex_2	BMDU - sex_2	Converged	Accepted AIC
Null		1	-275.98	553.96	NA	NA	NA	NA	NA	NA	yes	
Full		4	-112.78	233.56	NA	NA	NA	NA	NA	NA	yes	
Logistic	a	3	-113.51	233.02	3.44	2.92	4.02	3.44	2.92	4.02	yes	no
Probit		2	-122.32	248.64	3.03	0.00	3.54	3.03	0.00	3.54	yes	no
Log-logistic	b	4	-110.93	229.86	2.45	1.88	3.16	3.41	2.62	4.43	yes	no
Log-probit	b	4	-110.44	228.88	2.41	1.88	3.02	3.30	2.59	4.19	yes	yes
Weibull	b	4	-110.31	228.62	1.90	1.37	2.51	2.77	1.97	3.70	yes	yes
Gamma	b	4	-109.90	227.80	2.16	1.55	2.82	3.05	2.20	4.00	yes	yes
Two-stage	b	4	-110.90	229.80	1.75	1.14	2.57	2.62	1.69	3.81	yes	yes

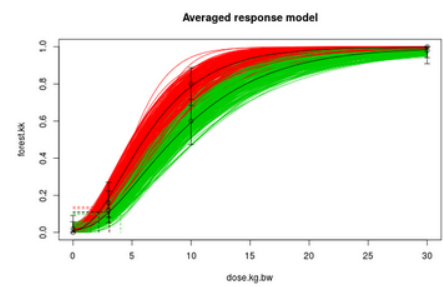
Download

Model Averaging

	Logistic	Probit	Log-logistic	Log-probit	Weibull	Gamma	Two-stage
Estimated model weights	0.02	0.00	0.12	0.19	0.22	0.33	0.12

	BMD	BMDL	BMDU
sex_1	2.17	1.51	3.08
sex_2	3.06	2.16	4.04

Download



Download

Figure 10: Quantal response data with covariate: fitted models results.

Model	Included covariate(s)	Number of parameters	Log-likelihood	AIC	BMD - sex_1	BMDL - sex_1	BMDU - sex_1	BMD - sex_2	BMDL - sex_2	BMDU - sex_2	Converged	Accepted AIC
Null		1	-275.98	553.96	NA	NA	NA	NA	NA	NA	yes	
Full		4	-112.78	233.56	NA	NA	NA	NA	NA	NA	yes	
Logistic		2	-117.75	239.50	3.35	2.84	3.94	3.35	2.84	3.94	yes	no
Logistic	a	3	-113.51	233.02	3.44	2.92	4.02	3.44	2.92	4.02	yes	no
Logistic	b	3	-113.66	233.32	2.88	2.36	3.49	3.93	3.27	4.68	yes	no
Logistic	a, b	4	-113.21	234.42	3.15	2.44	4.00	3.68	2.96	4.54	yes	no
Probit		2	-122.32	248.64	3.03	0.00	3.54	3.03	0.00	3.54	yes	no
Log-logistic		3	-113.59	233.18	2.80	2.21	3.51	2.80	2.21	3.51	yes	no
Log-logistic	a	4	-113.21	234.42	2.75	2.19	3.37	2.75	2.19	3.37	yes	no
Log-logistic	b	4	-110.93	229.86	2.45	1.88	3.16	3.41	2.62	4.43	yes	no
Log-logistic	a, b	5	-110.01	230.02	2.46	1.89	3.09	3.57	2.69	4.72	yes	no
Log-probit		3	-113.01	232.02	2.76	2.22	3.34	2.76	2.22	3.34	yes	no
Log-probit	a	4	-112.61	233.22	2.72	2.20	3.25	2.72	2.20	3.25	yes	no
Log-probit	b	4	-110.44	228.88	2.41	1.88	3.02	3.30	2.59	4.19	yes	yes
Log-probit	a, b	5	-109.52	229.04	2.42	1.90	2.99	3.45	2.66	4.59	yes	yes
Weibull		3	-113.81	233.62	2.12	1.54	2.77	2.12	1.54	2.77	yes	no
Weibull	a	4	-113.35	234.70	2.10	1.53	2.73	2.10	1.53	2.73	yes	no
Weibull	b	4	-110.31	228.62	1.90	1.37	2.51	2.77	1.97	3.70	yes	yes
Weibull	a, b	5	-109.51	229.02	1.91	1.37	2.51	2.84	2.01	3.85	yes	yes
Gamma		3	-113.05	232.10	2.45	1.78	3.16	2.45	1.78	3.16	yes	no
Gamma	a	4	-112.61	233.22	2.42	1.76	3.08	2.42	1.76	3.08	yes	no
Gamma	b	4	-109.90	227.80	2.16	1.55	2.82	3.05	2.20	4.00	yes	yes
Gamma	a, b	5	-109.08	228.16	2.16	1.55	2.81	3.14	2.25	4.21	yes	yes
Two-stage		3	-115.06	236.12	1.82	1.22	2.78	1.82	1.22	2.78	yes	no
Two-stage	a	4	-114.59	237.18	1.80	1.21	2.72	1.80	1.21	2.72	yes	no
Two-stage	b	4	-110.90	229.80	1.75	1.14	2.57	2.62	1.69	3.81	yes	yes
Two-stage	a, b	5	-110.10	230.20	1.77	1.14	2.59	2.70	1.72	4.00	yes	no

Figure 11: Quantal response data with covariate: fitted models results for all covariate combinations.

The image shows a software interface with two main panels. The left panel, titled 'Define variables', contains several input fields: 'Independent variable (e.g. dose)' set to 'ppm', 'Response variable(s)' set to 'mean', 'Variation statistic' set to 'sd', 'Type of variation statistic' with 'standard deviations' selected, 'Sample size for response variable:' with 'number' selected, and 'Models to fit' set to 'Set of models'. The right panel, titled 'Additional settings', has 'Value for CES' set to 0.05 and 'Confidence level for the BMD confidence intervals' set to 0.9. A small text box below the CES field explains that CES/BMR is the Critical Effect Size, defined as the estimated difference in response compared with the background response.

Figure 12: Continuous summary data: input values.

no best (accepted) model, a plot of the full model is shown. There we observe two groups of mean response per dose, so we suggest to perform an analysis with sex included as covariate.

When selecting the covariate ‘sex’, it is automatically tested for the background response parameter (a) and the potency parameter (BMD/CED). Retaining the previous settings for the other input fields, we obtain the results in Figure 14. Now, the models fit well to the data and AIC values are much lower than for the models without the covariate (see Figure 14). Notice that in the plot for the best model (Hill model 3) multiple lines are drawn. Each line corresponds with a given level of sex, from the data we know that the lowest, red curve corresponds with level 2.

Fit Model(s)

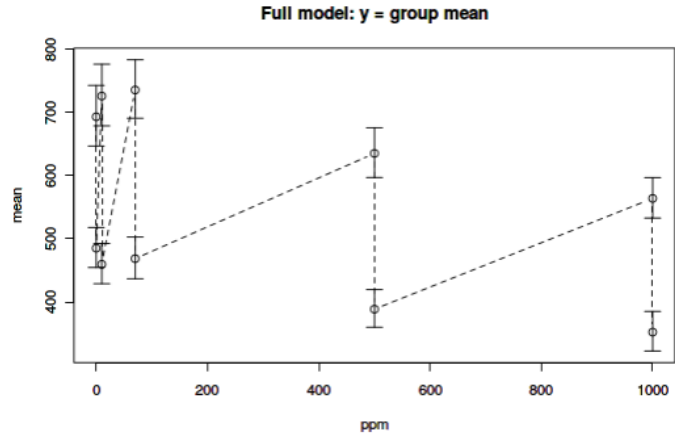
Response: mean

Fitted Models

Model	Included covariate(s)	Number of parameters	Log-likelihood	AIC	BMD	BMDL	BMDU	Converged	Accepted AIC
Null		2	-79.79	163.58	NA	NA	NA	yes	
Full		11	67.63	-113.26	NA	NA	NA	yes	
Exp model 3		4	-70.03	148.06	274.42	49.26	607.17	yes	no
Exp model 5		5	-69.45	148.90	439.77	68.57	524.18	yes	no
Hill model 3		4	-69.99	147.98	274.75	52.04	600.49	yes	no
Hill model 5		5	-69.45	148.90	423.39	79.14	523.38	yes	no

Download

Please consult a BMD specialist: None of the fitted models is at least as good as the full model: All fitted models' AIC values are larger than full model's AIC + 2



Download

Figure 13: Continuous summary data: fitted models results.

Fit Model(s)

Response: mean

Fitted Models

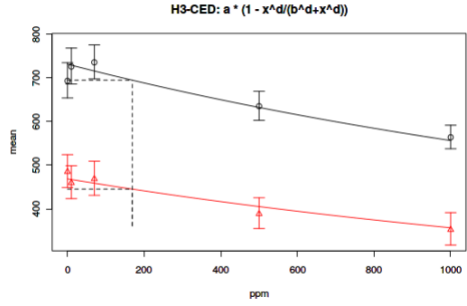
Model	Included covariate(s)	Number of parameters	Log-likelihood	AIC	BMD - sex_1	BMDL - sex_1	BMDU - sex_1	BMD - sex_2	BMDL - sex_2	BMDU - sex_2	Converged	Accepted AIC
Null		2	-79.79	163.58	NA	NA	NA	NA	NA	NA	yes	
Full		11	67.63	-113.26	NA	NA	NA	NA	NA	NA	yes	
Exp model 3	a	5	64.36	-118.72	162.29	58.55	309.57	162.29	58.55	309.57	yes	yes
Exp model 5	a	6	65.22	-118.44	424.27	306.92	457.85	424.27	306.92	457.85	yes	yes
Hill model 3	a	5	64.47	-118.94	169.00	64.27	312.76	169.00	64.27	312.76	yes	yes
Hill model 5	a	6	65.22	-118.44	406.72	118.94	449.30	406.72	118.94	449.30	yes	yes

Download

Lowest BMDL(s): 58.55; 45.44

Highest BMDU(s): 486.02; 482.67

Best model(s): Hill model 3 with covariate for a



Download

Figure 14: Continuous summary data with covariate: fitted models results.