User Manual and Technical Guidance
for
the Bayesian Benchmark Dose (BBMD) Analysis System

Version 1.1.0.0 (2016.10.27)

This document primarily serves as a user manual to provide introductions for using this system properly. In the second part of this document, a quick technical guidance is provided to help users better understand the Bayesian statistical methodology employed in this computational system and therefore assist users interpret the results.

Contents

I. Part I. User Manual

I.1 Access the Online System

The URL of the BBMD online system is https://benchmarkdose.org. Google Chrome and Firefox are two recommended web-browsers for using the system. The front page of the BBMD system is displayed in Figure 1 below.

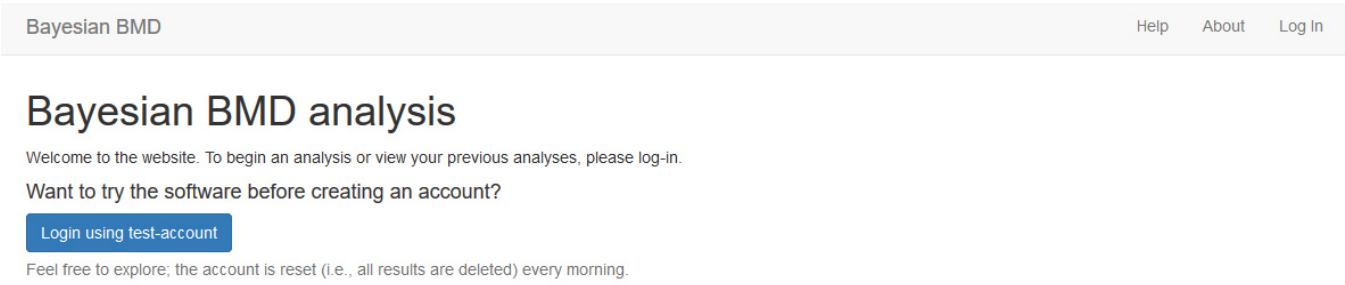


Figure 1. Front Page of the BBMD System

I.2 Login the System and Create a User Account

To use the system for BMD analysis, users need to login the system first by clicking "Log in" on the top right corner (shown in Figure 1). The "Log-in" page is shown in Figure 2 below. Email address (i.e., the username) and password are required to Log-in an account.

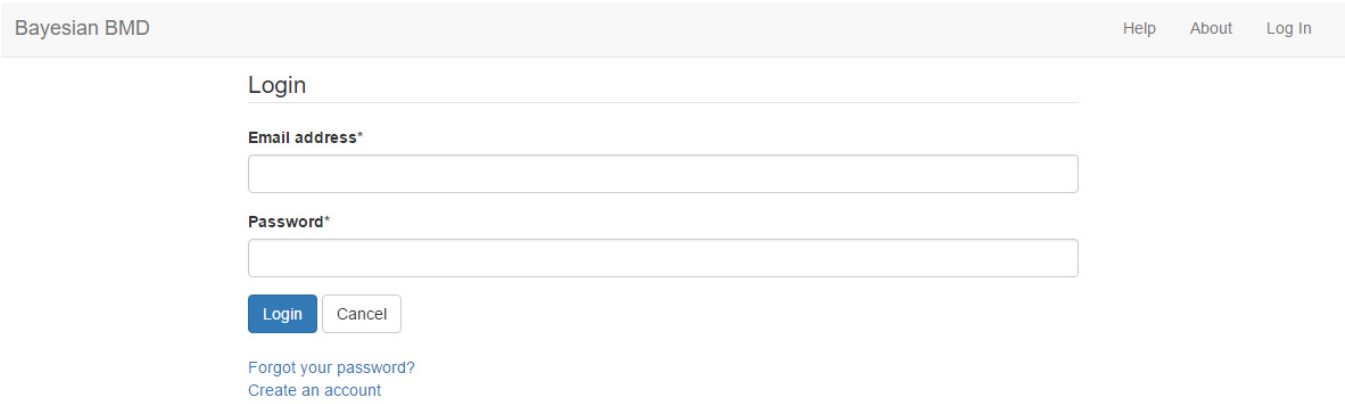


Figure 2. Login Page of the BBMD System

If this is the first time using this system, you need to create an account by clicking the "Create an account" link at the botton of the login page shown in Figure 2.

The “Create an account” page is shown in the Figure 3. Just like most online systems, email address (used as an account name) and entering password twice are needed to create an account. The main benefit of creating a personal account is that your previous analyses will be saved in your account for review or update.

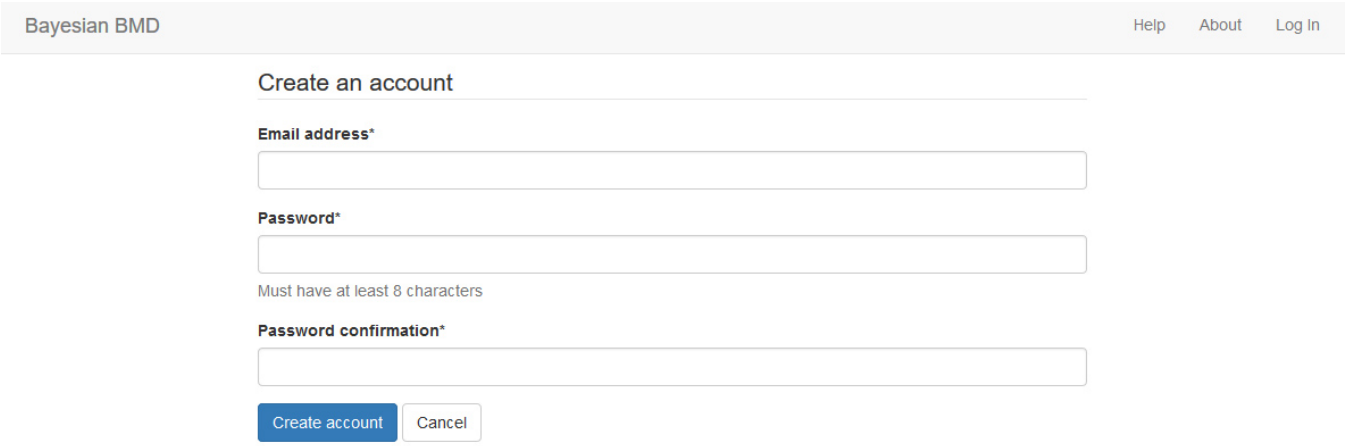


Figure 3. “Create an account” Page of the BBMD System

Currently, at the testing stage, the system also provides a public account for new users to test the BBMD system. If logged in the test account, user can use this account to test all the functions except that the results will be deleted periodically. The analyses results may also be viewed, modified or deleted by other users of the public account. The way to use the testing account is simply clicking the “Login using test-account” on the front page shown in Figure 1.

I.3 Start a New Analysis or Update an Existing Analysis

Once logged in, you now reach the page with options to access previous analyses (if you have any) or to start a new analysis by pressing “New run” on the top right corner, as shown in Figure 4.

Figure 4. The Summary Page of Existing Runs

If you select one of your previous analyses (e.g., “Example 2” or “Example 1” in Figure 4), you can review the results in this previous analysis. If “New run” is clicked, you will start a new analysis from scratch, which will be introduced In Section I.5.

I.4 Reviewing an Existing Analysis and Exporting Results

Click one of the names of the existing runs shown in Figure 4, results from the selected study will be accessed by the system and displayed. Figure 5 shows the review page of an existing analysis.

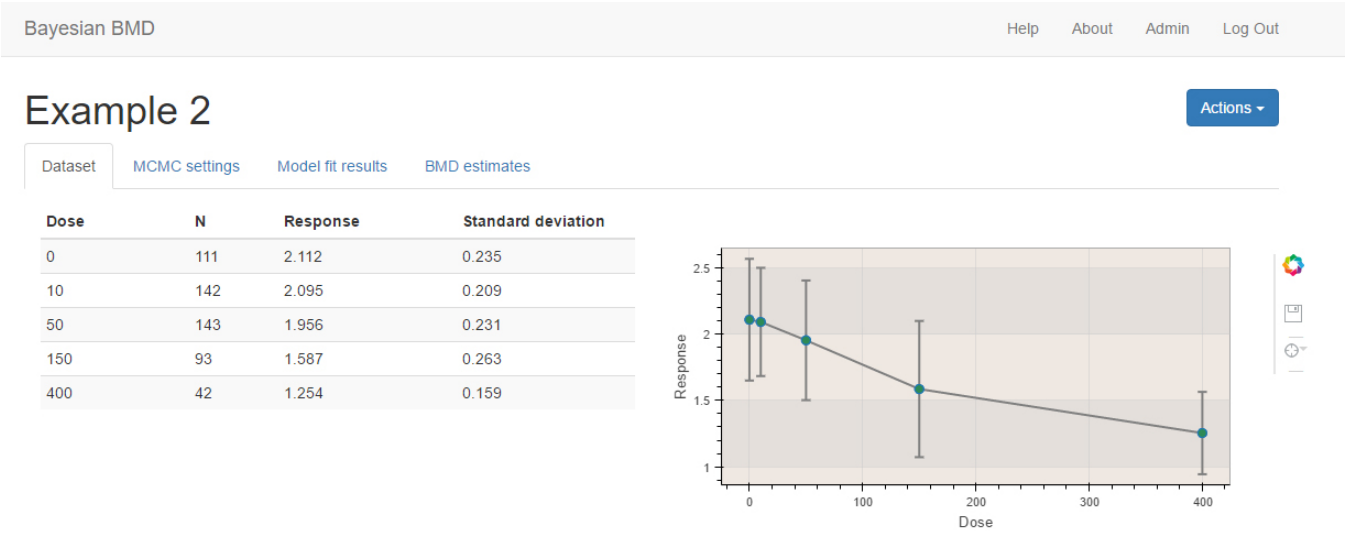


Figure 5. Reviewing an Existing Analysis

There are four tabs available for each of the analysis: “Dataset”, “MCMC settings”, “Model fit results”, “BMD estimates”. Clicking each of the tabs, the data or setting stored in this analysis will be shown. One or more tabs may be empty which indicates that the analysis was ended previously before the corresponding section had been finished. The detailed information regarding the contents in each of the tabs will be introduced in the following sections.

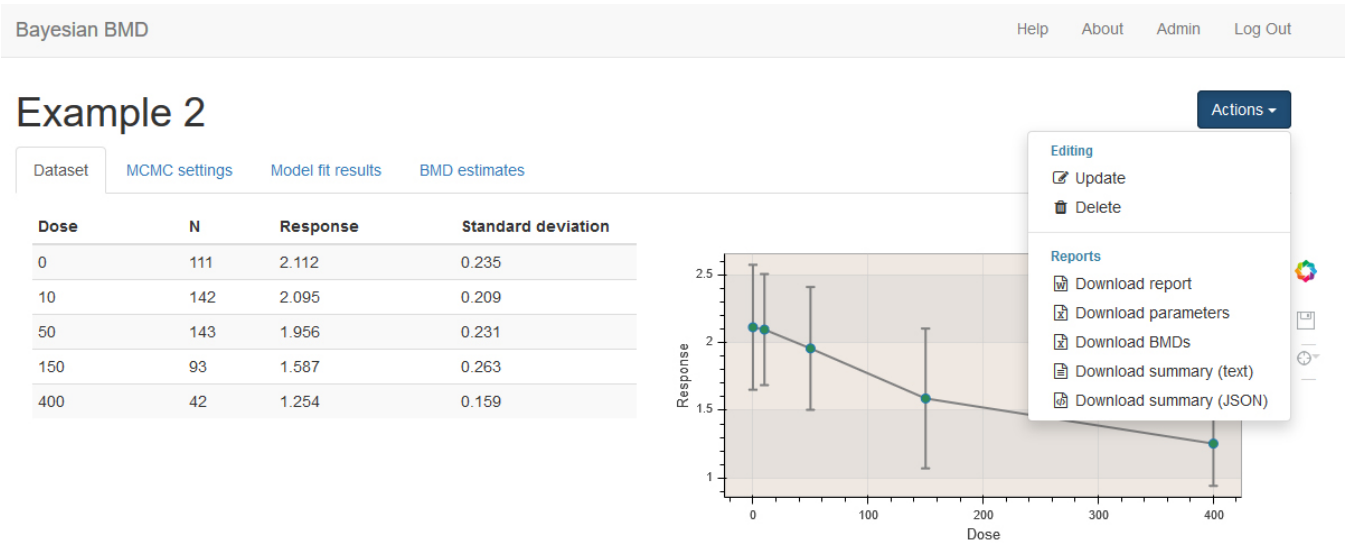


Figure 6. The Actions Available to Users for an Existing Analysis

Clicking the blue "Action" button on the top right corner, a list of action items will be available as shown in Figure 6. There are two categories of action items: "editing" and "reports". The function of each of the action items is described below:

- Update: this existing analysis can be modified by user to update the results
- Delete: this existing analysis can be deleted from user's account
- Download report: a summary report in Microsoft Word format containing all important information regarding the analysis will be downloaded
- Download parameters: a Microsoft Excel file containing posterior sample of all model parameters of all models included in the analysis will be downloaded
- Download BMDs: a Microsoft Excel file containing posterior sample of different BMDs estimated from all models included in the analysis will be downloaded
- Download summary (text): a summary report in Text format will be downloaded
- Download summary (JSON): a summary report in JSON format will be downloaded

Please note that the **reports and analysis data can be only downloaded on this review page, and an analysis can only be deleted under this condition**.

I.5 Creating a New Analysis or Updating an Existing Analysis

Once logged in a personal account (or the testing account), click "New Run" on the top right corner or choose the "update" option in the pull-down menu in review page for an existing analysis like what has been shown in Figure 6, then you reach the editing phase of an analysis session where you need to finish settings for at least three tabs as shown in Figure 7.

An automatically generated name, "New run Data, Time", is assigned to the newly started analysis. You can click "Edit name" button next to the analysis name as shown in the Figure 7, and then modify it to a more identifiable name. Without putting in any data, none of the tabs on the data analysis page is highlighted and accessible. Therefore, the first thing in an analysis is to input dose-response data.

I.5.1 Data Input

In the current version, the system is capable of analyzing the following four types of dose-response data. Use the pull-down menu called “Dataset type” to select one of the following data types:

a) Dichotomous summary: this type requires three values for each dose group (i.e., each input row): dose level, total number of subjects in that group, and the number of subjects affected
b) Dichotomous individual: this type requires two values for each input row (representing each subject): dose level, and “0” or “1” indicating that the subject is non-affected or affected respectively.
c) Continuous summary: this type requires four columns of data for each dose group: dose level, number of subjects in the dose group, mean value of the response, and standard deviation or standard error of the response.
d) Continuous individual: this type requires two values for each input row representing each individual subject: dose level and response

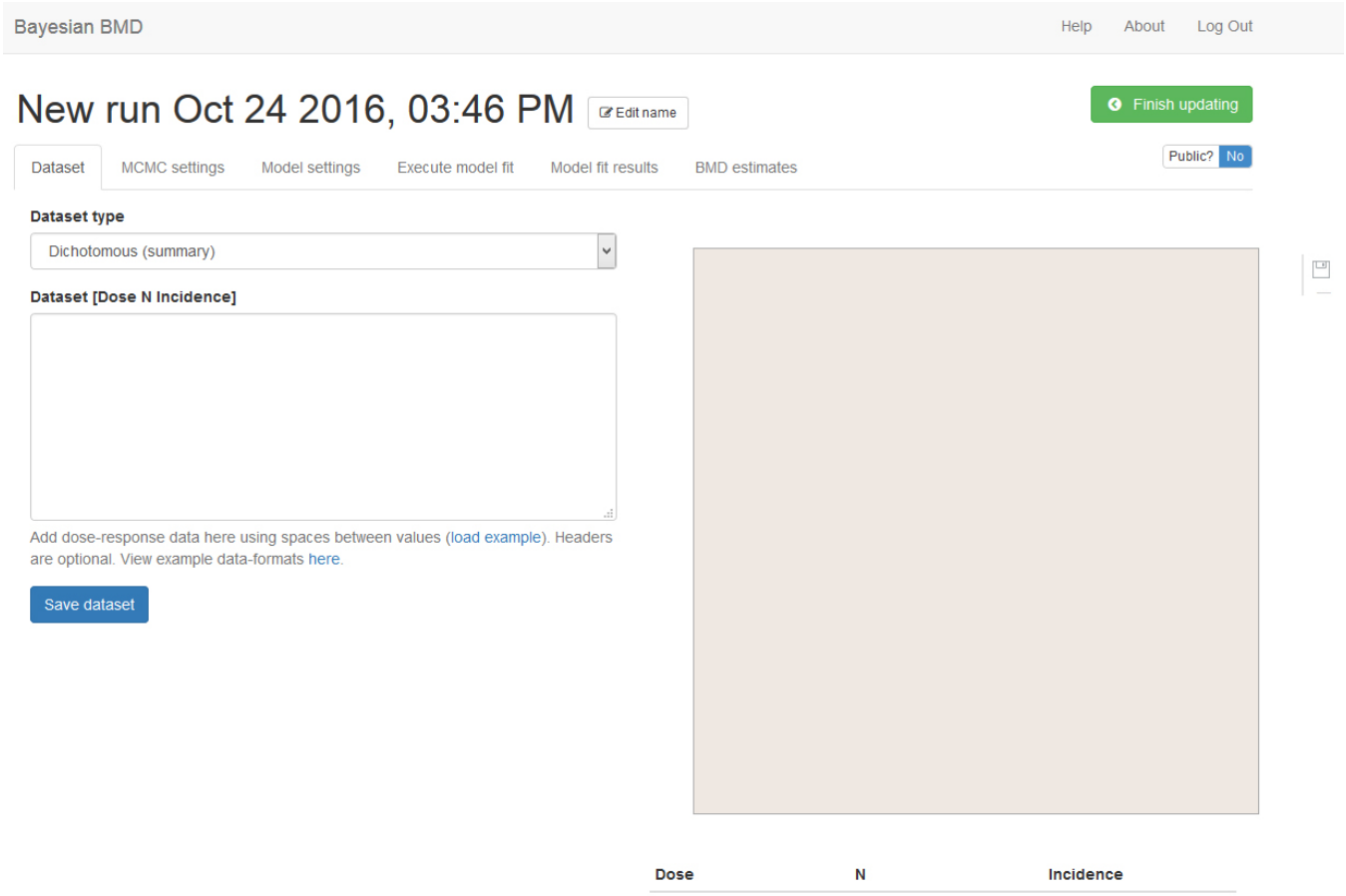


Figure 7. The Start Page of a New Analysis

Once you choose “Dichotomous summary” for dichotomous summary data, three columns (from left to right) are required for input: dose level, total number of subjects and number of subjects affected. The

values can be pasted or manually typed, using spaces between values. Different dose groups should be entered in different rows. Below is an example dichotomous summary dataset:

(Dose)	(# of Subjects)	(# of Affected)
0	50	1
15.5	49	4
30	50	8
50.6	48	21

If you choose “Dichotomous individual”, two columns are required. The two columns from left to right are: dose and indicator (either “0” representing no effect or “1” representing with effect). Each row is used for each individual subject. An example of dichotomous individual data is shown below:

(Dose)	Indicator (0 or 1)
0	0
0	0
0	0
0	1
0	0
10	0
10	1
10	0
10	1
10	0
25	1
25	0
25	1
25	1
25	0
50	1
50	1
50	1
50	1

For “Continuous summary” data type, four columns are needed to describe each dose group, which are dose, number of subjects, the mean value of response, and the standard deviation of the response (or the standard error of the response). A pull-down menu is used to specify what the last column is. The dataset below is an example of the continuous summary data:

(Dose)	# of Subjects	Mean	Std Dev
0	10	2.82	0.17
100	10	2.91	0.16
200	10	2.95	0.2
400	9	3.22	0.25

The last data type is “Continuous individual” which only requires two columns (from left to right): dose level and response. The table below shows you an example of this data type.

(Dose)	(Response)
0	351.3
0	350.3
0	359.8
0	360.7
0	357.4
2.5	349.8
2.5	352.1
2.5	346.3
2.5	344.7
2.5	350.1
5	340.2
5	341.1
5	345.5
5	331.9
5	347.4
20	331.1
20	320.9
20	319.4
20	308.9
20	314.3

Once the data have been entered, press “Save dataset” to save the dataset. Please refresh your browser to make sure data set has been successfully saved. Once the data set is successfully saved, the data will be visually displayed and summarized in a table as shown in the Figure 8. Additionally, the tabs of “MCMC settings”, “Model settings” and “Execute model fit” are now highlighted.

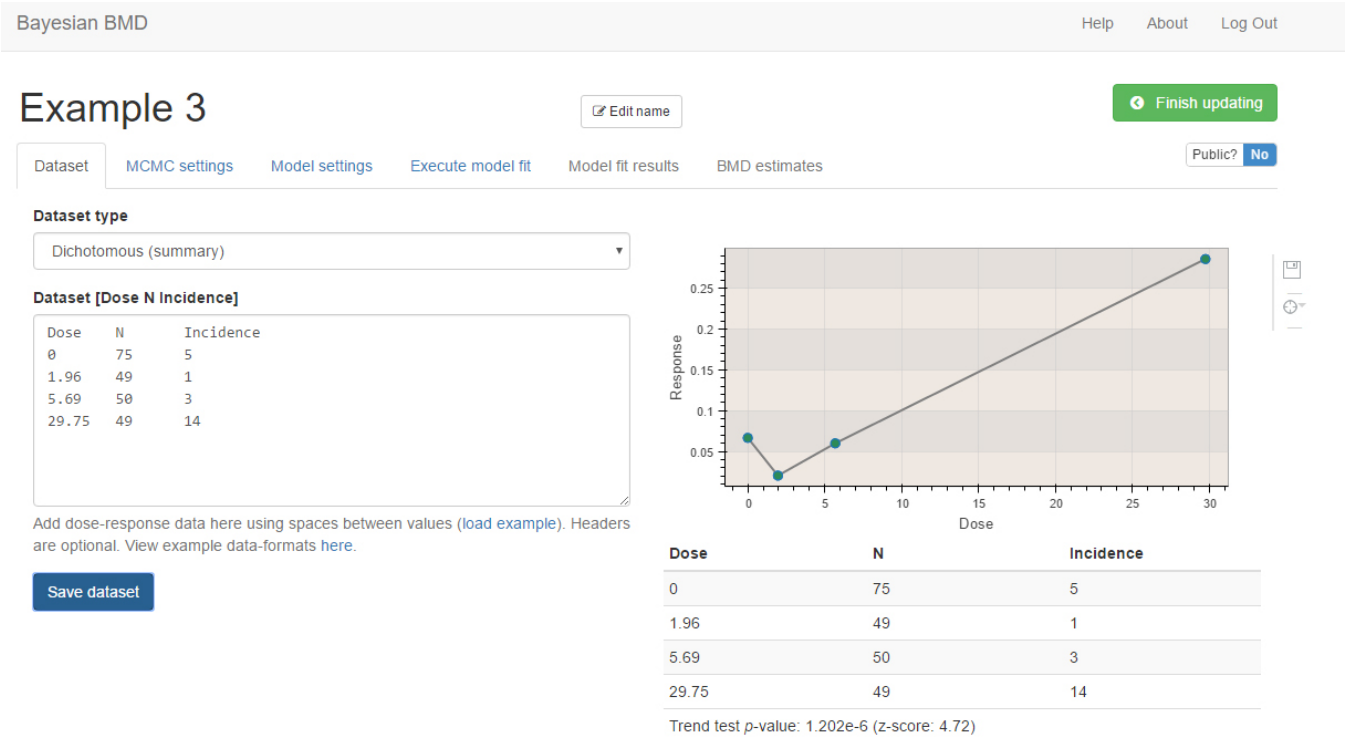


Figure 8. Completed Data Input

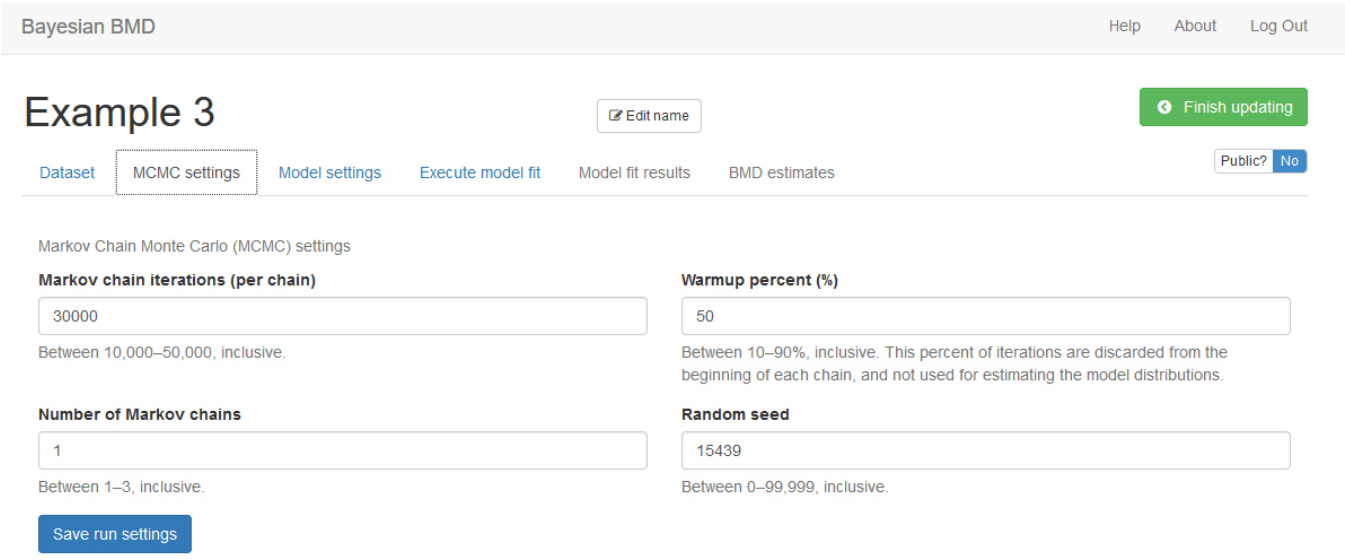


Figure 9. MCMC settings

I.5.2 MCMC Settings

On this tab shown in Figure 9, you need to specify some settings for the MCMC algorithms.

"Iterations" is the length of MCMC chain, i.e., the number of posterior sample in each MCMC chain. Default value is 30,000. The allowable range is any integer between 10,000 and 50,000.

"Number of chains" is the number of Markov Chain to be sampled. Default value is 1. Allowable range is 1 ~ 3.

"Warmup percent (%)", the percent of sample in each Markov Chain will be discarded from the final posterior sample. Default value is 50% with an allowable range of 10% ~ 90%.

So, using the default values, the final number of posterior sample (without the warmup sample) you can get is:

$$30000 \times 1 \times (1 - 50\%) = 15000$$

"Seed" is random seed number used in the MCMC algorithms. The number is randomly generated, but you can specify the number for the purpose of reproduction.

Once these values are specified, click "Save run settings" to save the MCMC settings. Default settings are generally acceptable. However, results in the next step will provide important information that can help you judge if the MCMC settings are appropriate. Based on our testing, the **default settings are adequate for most of the commonly seen dose-response shapes, so it is suggested that using the default settings for your initial run**.

I.5.3 Model Settings

Once data input and MCMC settings are completed, go to the next tab "Model settings". In this step, you should choose the model(s) to fit the data. Figure 10 shows how this page looks like after click the "Add new model" link on the left panel.

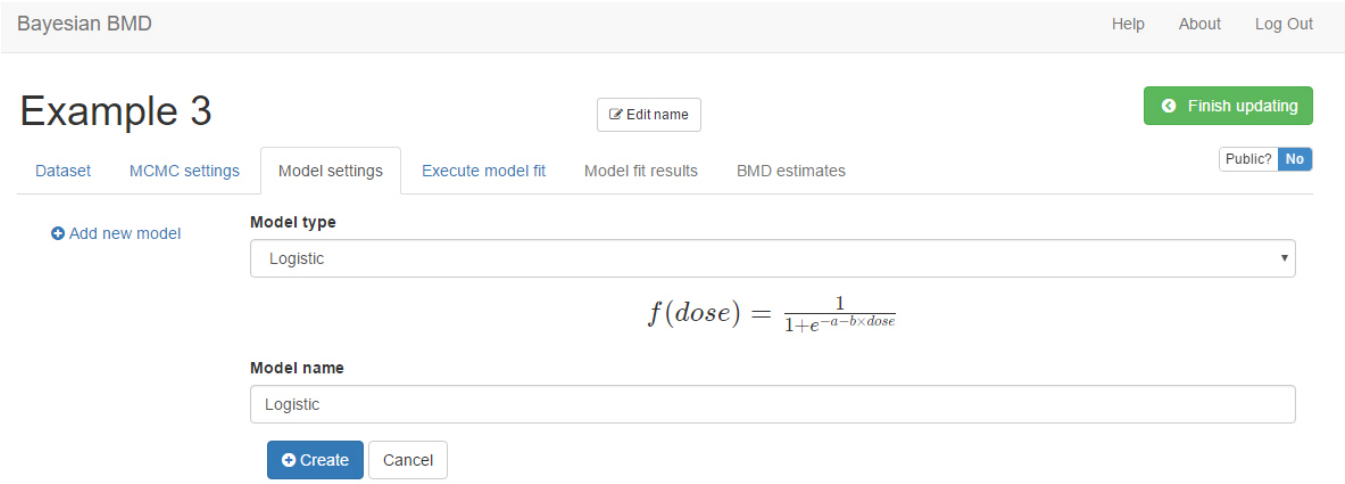


Figure 10. The Page of "Model Settings"

A. Add a dose-response model

Three steps are required to add a model to the list of models to be analyzed:

1) Click "Add new model" link on the left panel
2) Select one model from the pull-down menu, then give an identifiable name to the model. For models with a power parameter, you need to choose a restriction value for the power parameter. There are three options available in the current system: 0, 0.5, and 1.
3) Press "Create" button to add the model to the list of models on the left panel.

To add another model, repeat the steps 1) to 3). A same model with different settings (e.g., the restriction value put on the power parameter) can be added twice as two different models.

B. Update or delete a model from a the model list

As shown in Figure 11, you can update or delete a model which is already in the list of models by following the following three steps:

1) Click the name of the model you want to update or delete in the list of models on the left panel
2) If you want to change the current settings of the model, modify the setting option accordingly, including the model type, model name, and the restriction value
3) Click "Update" below the settings to update the setting, or "Delete" to delete the model

The models shown on the left panel are the models will be analyzed by the system. The available models for continuous and dichotomous data are listed in the Table I below.

Table I. Available dose-response models in the BBMD system

Dose-Response Models for Dichotomous Data	Dose-Response Models for Continuous Data
Quantal-linear model: $f(d) = a + (1-a)\times[1-\exp(-b\times d)]$, $0 \le a \le 1,\ \ b \ge 0$	Linear model: $f(d) = a + b\times d,\ \ a > 0$
Probit model: $f(d) = \Phi\,(a + b\times d),\ \ b \ge 0$	Power model: $f(d) = a + b\times d^{g},\ \ a > 0,\ \ g \ge restriction$
Logistic model: $f(d) = \frac{1}{1+\exp\,(-a-b\times d)},\ \ b \ge 0$	Michaelis-Menten model: $f(d) = a + \frac{b\times d}{c+d},\ \ a > 0,\ \ c > 0$
Weibull model: $f(d) = a + (1-a)\times[1-\exp(-b\times d^{g})]$, $0 \le a \le 1,\ \ b \ge 0,\ \ g \ge restriction$	Hill model: $f(d) = a + \frac{b\times d^{g}}{c+d^{g}}$, $a > 0,\ \ c > 0, g \ge restriction$
Multistage (2nd degree) model: $f(d) = a + (1-a)\times[1-\exp(-b\times d - c\times d^{2})]$, $0 \le a \le 1,\ \ b \ge 0\,, c \ge 0$	Exponential model 2: $f(dose) = a\times\exp(b\times d)\,,\ \ a > 0$
LogLogistic model: $f(d) = a + \frac{1-a}{1+\exp\,[-b-g\times\log\,(d)]}$, $0 \le a \le 1,\ \ g \ge restriction$	Exponential model 3: $f(dose) = a\times\exp(b\times d^{g})\,,$ $a > 0,\ \ g \ge restriction$
LogProbit model: $f(d) = a + (1-a)\times\Phi\,[b + g\times\log(d)]$ $0 \le a \le 1,\ \ g \ge restriction$	Exponential model 4: $f(dose) = a\times[c-(c-1)\times\exp(-b\times d)]$, $a > 0,\ \ b > 0,\ \ c > 0$

Dichotomous Hill model: $$f(d) = a\times b + \frac{a - a\times b}{1 + \exp\left[-c - g\times\log(d)\right]},$$ $$0 < a \leq 1, 0 < b < 1,\ \ g \geq restriction$$	Exponential model 5: $$f(dose) = a\times[c - (c-1)\times\exp(-(b\times d)^{g})],$$ $$a > 0,\ \ b > 0,\ \ c > 0,\ \ g \geq restriction$$

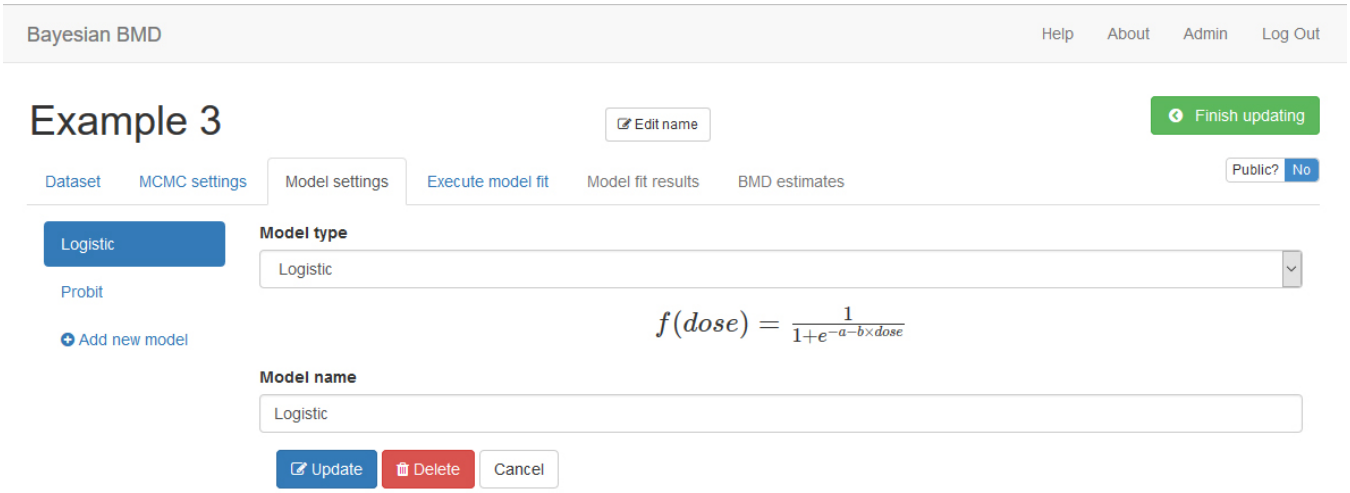


Figure 11. Update or Delete a Model in the List

I.5.4 Execute model fit

Once the data input, MCMC settings, and model settings are completed, it's ready to execute the MCMC simulation for model fitting on the "Execute model fit" tab as shown in Figure 12. If the settings in first three tabs have been successfully saved, all three check marks will be green indicating that the analysis can be executed. If one or more check marks are red, you need to go back to the corresponding tab and save the settings.

After clicking the "Execute" button, the system will give an update regarding the time that has been used for this analysis below the "Execute" button. For an analysis containing all eight dose-response model using the default setting, it takes about 30 seconds to 1 minute to finish the analysis. After the running is finished, the fifth and sixth tab "Model fit results" and "BMD estimates" previously greyed out will become available.

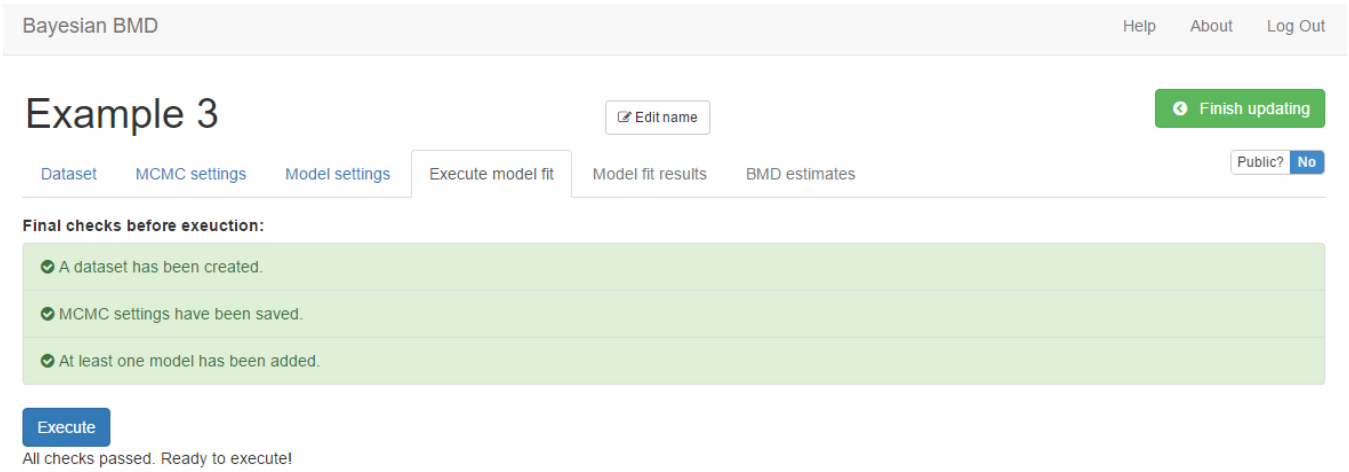


Figure 12. The Page of Execute Model Fit

I.6 Review Model Fit Results

On the "Model fit results" tab, the model fitting results obtained from the previous step are displayed. Click the name of one of the models on the left panel, then the results will be shown on the right as shown in Figure 13, including the textual output of model parameter estimation, dynamic dose-response plot, posterior predictive p-value, model weight, correlation matrix, and graphical output of posterior sample of the model parameters (hidden by default).

When you move mouse on the dose-response curve, the response with its 90th percentile interval at this particular dose level will be shown on the plot. Click the "Show parameter charts", two plots (posterior sample trace plot and estimated probability density plot) will be displayed for each of the parameters in this dose-response model.

Basically, this is a results display tab, meaning that you can only review the results but cannot give the system additional inputs to modify the results.

The methodology used to generate these output values as well as the interpretation of these quantities will be introduced in the Technical Guidance Section.

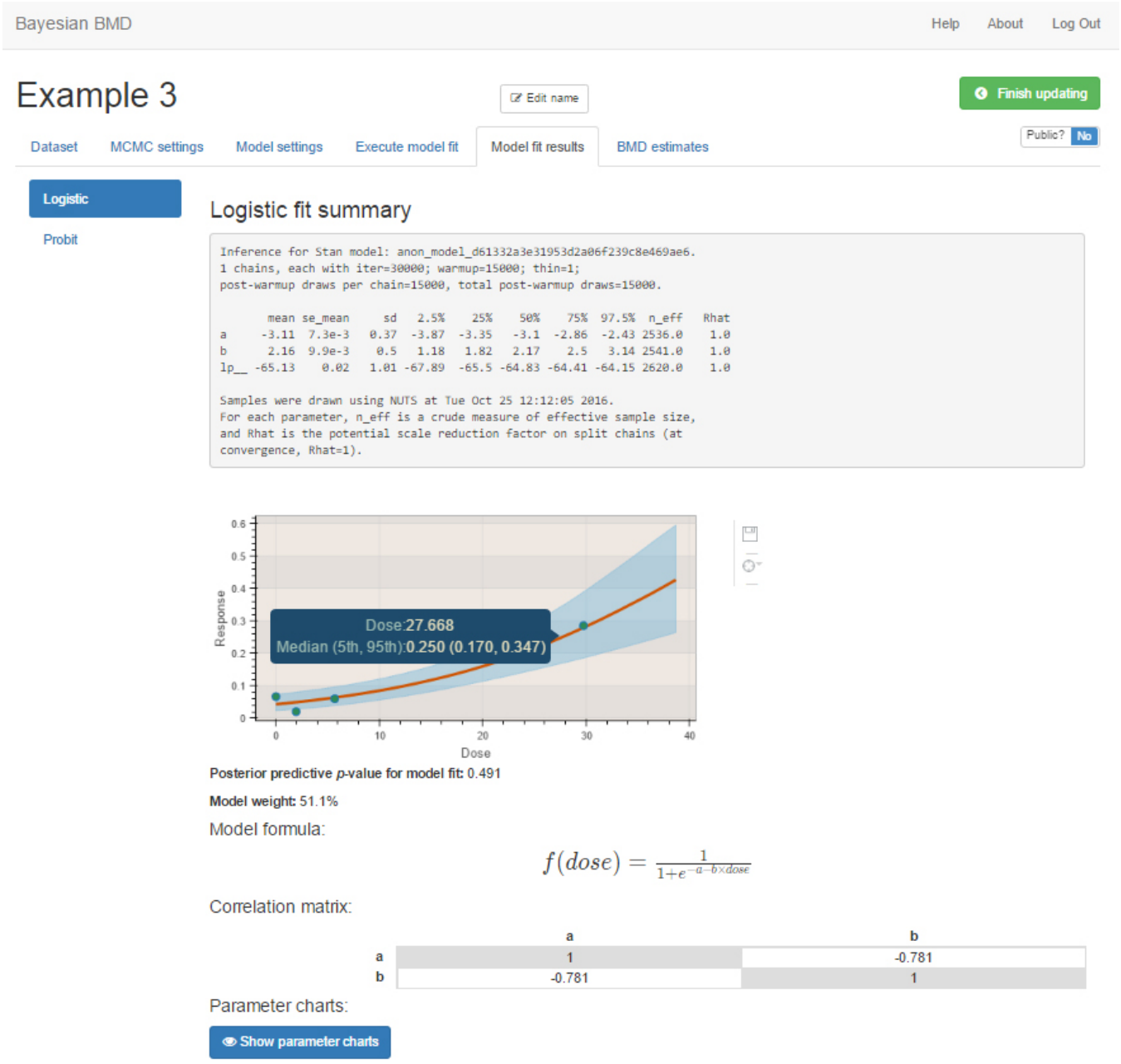


Figure 13. Results Shown on the "Model fit Results" Page

I.7 Benchmark Dose Analysis

On this page, you can calculate the BMD estimates of your interest. The settings for BMD calculation are slightly different between the analysis for dichotomous data and continuous data, therefore, they will be introduced separately.

I.7.1 BMD Analysis for Dichotomous Data

Figure 14 is a screenshot of the "BMD estimates" tab. To create a BMD analysis, you need to follow the four steps below:

(1) Name the BMD analysis using an easily identifiable name in the "Model name" box;
(2) Specify a BMR value in the "Benchmark response value" box;
(3) Give prior model weight to the models included in this analysis. Giving 0 weight to a particular model can exclude the model from model-averaged BMD calculation. The sum of the weights assigned to the individual models are not necessarily required to be 1. The system will automatically convert the the. For example, if you give 1 to Logistic model and 3 to Probit model, the system will convert these value to 25% prior weight for the Logistic model and 75% prior weight for the Probit model;
(4) Click the "Create" button to execute the BMD analysis using the settings just specified

Once the BMD analysis is successfully created, the name of the analysis will show on the left panel and the results will be shown on the right panel as shown in Figure 15. Similarly as the procedure to add, edit and delete a dose-response model in the "Model settings" the tab, BMD analyses can be added, edited or deleted. After clicking the name of an existing analysis, we can update or delete the analysis using the buttons shown at the bottom of Figure 15. To add a new BMD analysis, just click the blue "Add new BMD" button on the left panel and then repeat the steps (1) to (4) above.

For dichotomous data, the BMD will be calculated for both the added BMR definition and extra BMR definition. The BMDs are defined by the following equations:

Added risk: $f(BMD) - f(0) = BMR$ Extra risk: $\frac{f(BMD)-f(0)}{1-f(0)} = BMR$

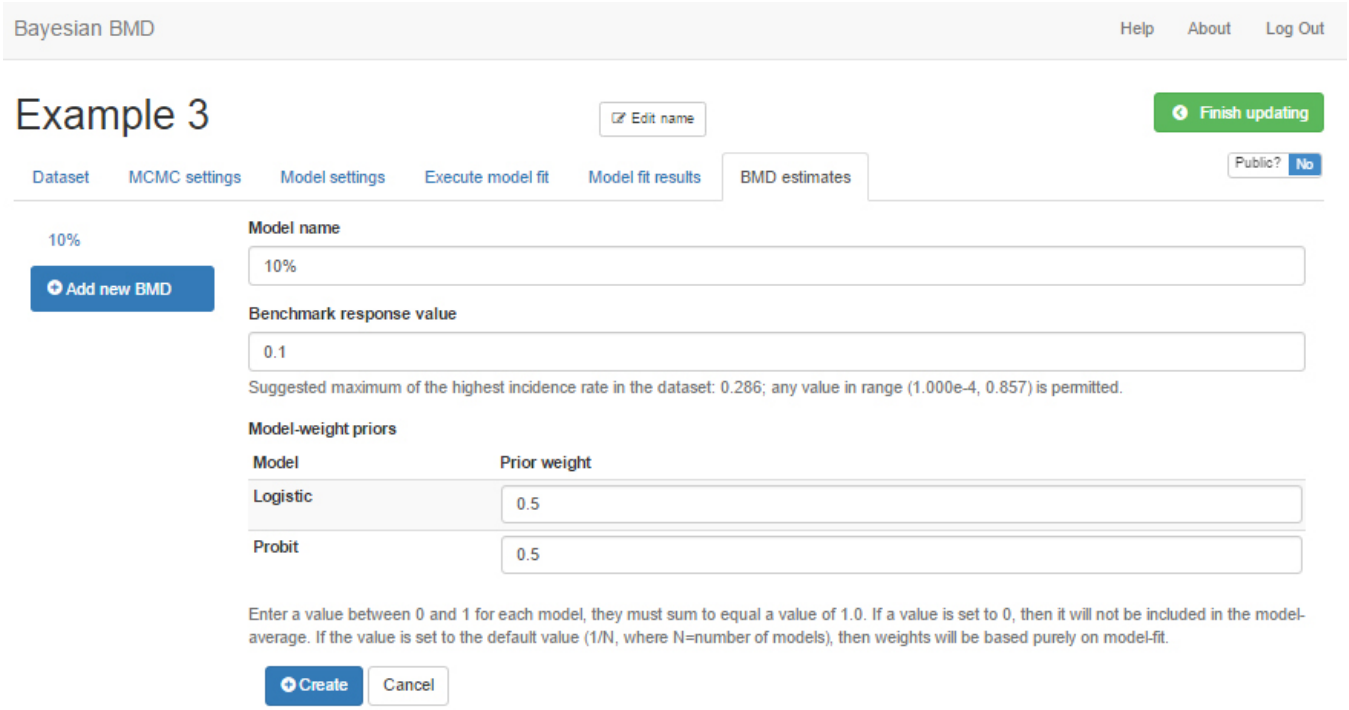


Figure 14. BMD Analysis Input for Dichotomous Data

Bayesian BMD

Help About Log Out

Example 3

Edit name

Finish updating

Public? No

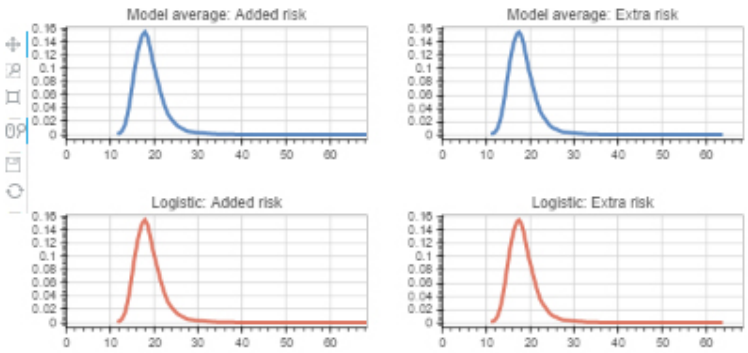
Dataset MCMC settings Model settings Execute model fit Model fit results BMD estimates

10%

Add new BMD

10%

BMD estimates



BMD summary table

Added risk

Statistic	Model average	Logistic
Prior model weight	N/A	0.500
Posterior model weight	N/A	0.511
BMD (median)	**18.0**	**18.3**
BMDL (5%)	**15.2**	**14.7**
25%	16.7	16.7
Mean (SD)	18.4 (2.56)	18.9 (3.49)
75%	19.5	20.3
95%	22.7	24.7

Extra risk

Statistic	Model average	Logistic
Prior model weight	N/A	0.500
Posterior model weight	N/A	0.511
BMD (median)	**17.6**	**17.8**
BMDL (5%)	**14.7**	**14.2**
25%	16.2	16.2
Mean (SD)	17.8 (2.44)	18.3 (3.32)
75%	18.9	19.8
95%	21.9	23.8

Model name

10%

Benchmark response value

0.1

Suggested maximum of the highest incidence rate in the dataset: 0.286; any value in range (1.000e-4, 0.857) is permitted.

Model-weight priors

Model	Prior weight
Logistic	0.5

Enter a value between 0 and 1 for each model, they must sum to equal a value of 1.0. If a value is set to 0, then it will not be included in the model-average. If the value is set to the default value (1/N, where N=number of models), then weights will be based purely on model-fit.

Update Delete Cancel

Figure 15. BMD Estimation Results Shown on the “BMD Estimates”

I.7.2 BMD Analysis for Continuous Data

For continuous data, the BBMD system provides two ways to define BMD: (1) based on central tendency and (2) based on tails (hybrid approach). The basic steps to create a BMD analysis for continuous data are almost identical to the procedure for dichotomous data, except the settings for the BMR.

“Central tendency” is the default option for BMD estimation method for continuous data. When this method is selected, an “Adversity measure” needs to be further specified as shown in Figure 16.

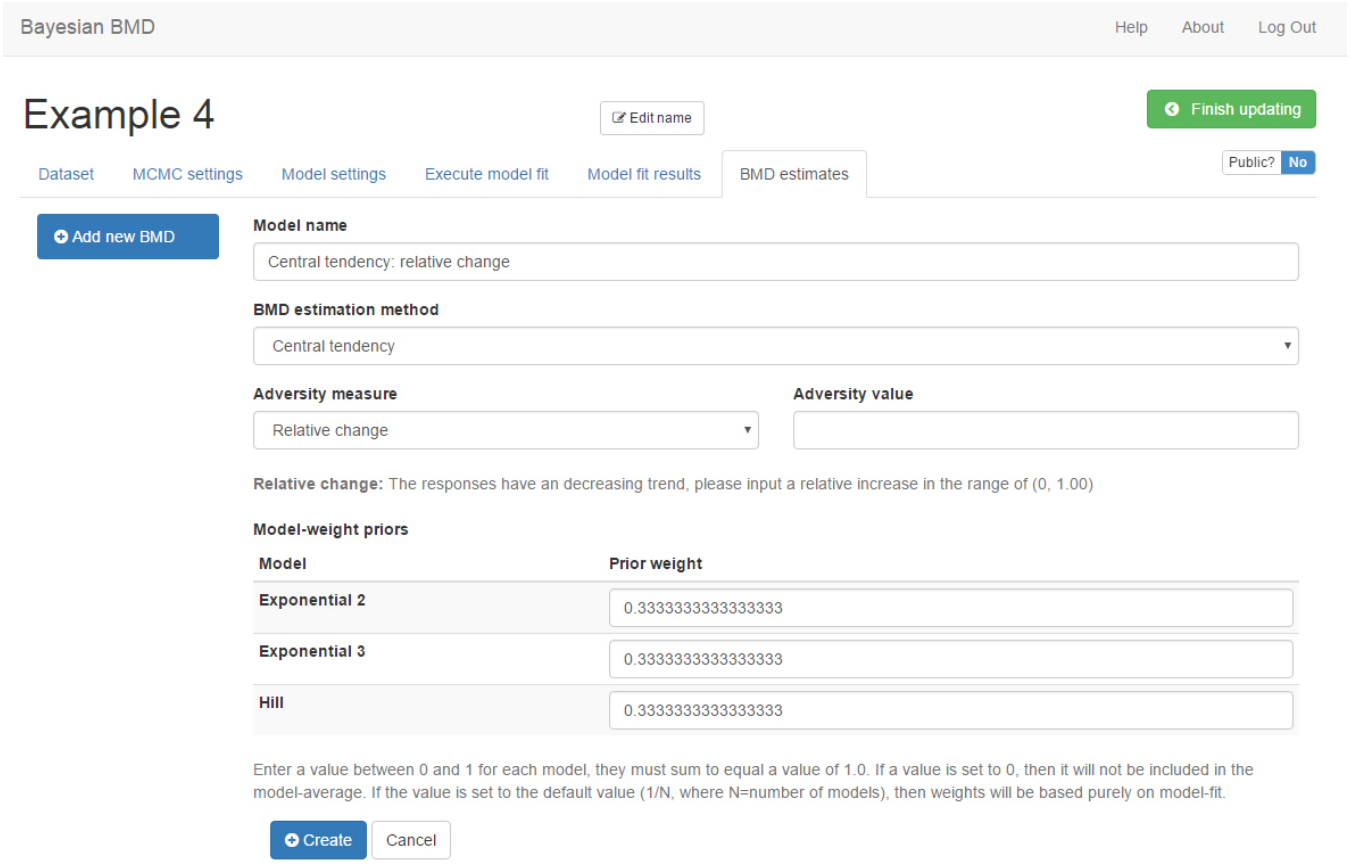


Figure 16. Settings for Continuous Data BMD Calculation Based on Central Tendency

There are three options to specify an adversity:

1) Relative change:
   For this option, you need to input a value of relative change, e.g., 20%. This means that if the central tendency changes 20% from the control, it will be considered as adverse and the BMD will be calculated accordingly, using the following equation:

$$f(BMD) \pm f(0) = Relative\ Change \times f(0)$$

The plus/minus sign on the left-hand side is related to the dose-response trend, if increasing, then it is "+", otherwise it is "-".

2) Absolute change:
For this option, you need to input a value of absolute change, e.g., 3.2. This means that if the central tendency changes 3.2 from the control, it will be considered as adverse and the BMD will be calculated accordingly, using the following equation:

$$f(0) \pm Absolute\ Change = f(BMD)$$

The plus/minus sign on the left-hand side is related to the dose-response trend, if increasing, then it is "+", otherwise it is "-".

3) Cutoff:
For this option, you need to input a value of cutoff, e.g., 22.5. This means that if the central tendency is equal to the cutoff value specified, it will be considered as adverse and the BMD will be calculated accordingly, using the following equation:

$$f(BMD) = cutoff$$

The allowable range for the values of these three options will be automatically calculated based on the trend of the dose-response data and shown.

If you choose to use the hybrid approach to estimate the BMD, you need to first specify an adversity value and then input a BMR value as shown in Figure 17. There are two ways to specify the adversity, either specifying a cutoff value ("Absolute cutoff value", i.e., depending on increasing or decreasing dose-response trend, above or below a value will be considered as adverse) or using a percentile value of the control ("Control group percentile", e.g., the below 1st percentile or above 99th percentile of the control distribution is considered as adverse depending on decreasing or increasing). Let's use $Q(0)$ to represent the quantile of the adversity value at control dose and use $Q(BMD)$ to represent the quantile of the adversity value at the BMD level, then the following equation should be satisfied:

For increasing trend

$$Q(0) - Q(BMD) = BMR \text{ for added risk}$$

$$\frac{Q(0)-Q(BMD)}{1-Q(0)} = BMR \text{ for extra risk}$$

For decreasing trend

$$Q(BMD) - Q(0) = BMR \text{ for added risk}$$

$$\frac{Q(BMD)-Q(0)}{Q(0)} = BMR \text{ for extra risk}$$

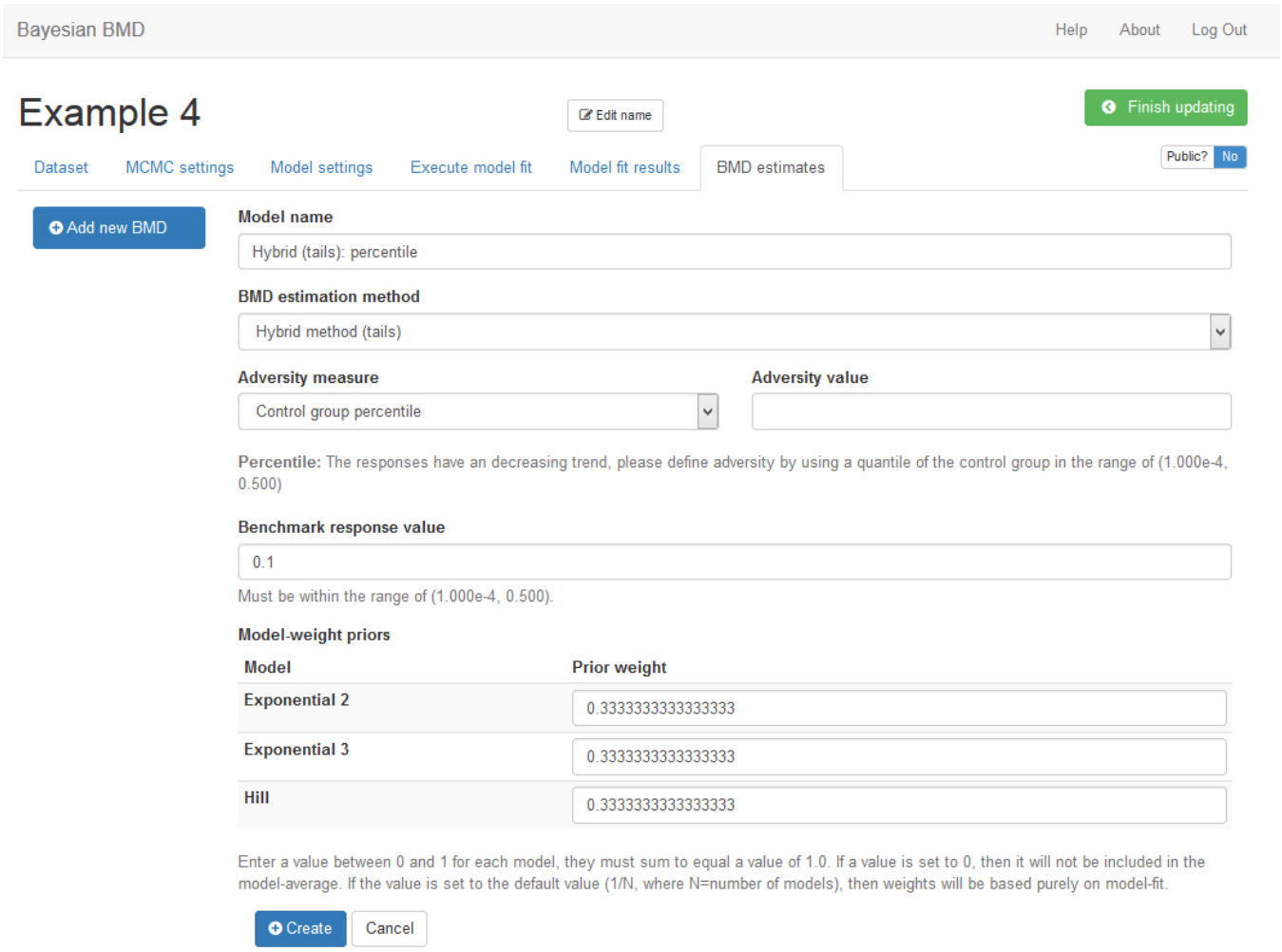


Figure 17. Settings for Continuous Data BMD Calculation Based on Tails

I.8 Change to Review Page, Share an Analysis, and Change to another Analysis

In the process of creating/updating a new analysis, you can stop editing the analysis and go back to the review page by click the green button “Finish updating” on the upper right corner.

By default, all analyses in a personal account can only be accessed by the owner of the account. However, an analysis can be shared with and viewed by other if the account owner change the default setting from private to public by clicking the “Public? No”. After it becomes “Yes Public?”, the owner can copy the URL and share it with others. Others can access and review (but cannot edit) this analysis by visiting the specific URL.

At any time during the updating or reviewing stage you want to change to another existing analysis, you can click “Bayesian BMD” button on the top left corner to switch to the summary page for the existing analyses and access another analysis.

II. Part II. Technical Guidance

II.1 Data Pre-Analysis

A trend test (using the Cochran-Armitage linear trend test, same as the BMDS) is performed once a dichotomous dataset has been successfully saved. A p-value and z-score will be reported below the data table in the "Data preview" section in the "Dataset" tab. The trend test results can be used to judge if the dichotomous dataset is appropriate for BMD modeling.

II.2 Model Settings

In this version of the BBMD system, uniform distribution is used as the prior distribution for all of the dose-response model parameters. The prior distributions for each model are listed below:

II.2.1 Dichotomous Data

For dichotomous data, there are eight models:

(1) Quantal-linear model (i.e., multistage-1st model)

$$f(dose) = a + (1-a)\times(1-e^{-b\times dose})$$

$$a \sim Uniform(0,1);\ b \sim Uniform(0,100)$$

(2) Logistic model

$$f(dose) = \frac{1}{1+e^{-a-b\times dose}}$$

$$a \sim Uniform(-50,50);\ b \sim Uniform(0,100)$$

(3) Probit model

$$f(dose) = \boldsymbol{\Phi}\,(a + b\times dose)$$

$$a \sim Uniform(-50,50);\ b \sim Uniform(0,100)$$

(4) Weibull model

$$f(dose) = a + (1-a)\times(1-e^{-c\times dose^{b}})$$

$$a \sim Uniform(0,1);\ b \sim Uniform(restriction,15);\ \ c \sim Uniform(0,50)$$

Where “restriction” is a user defined value and can be 0, 0.5 or 1.

(5) Multistage-2nd model

$$f(dose) = a + (1-a)\times(1 - e^{-b\times dose - c\times dose^2})$$

$$a\sim Uniform(0,1);\ b\sim Uniform(0,100);\ \ c\sim Uniform(0,100)$$

(6) LogLogistic model

$$f(dose) = a + \frac{1-a}{1+e^{-c-b\times \log(dose)}}$$

$$a\sim Uniform(0,1);\ b\sim Uniform(restriction,15);\ \ c\sim Uniform(-5,15)$$

Where “restriction” is a user defined value and can be 0, 0.5 or 1.

(7) LogProbit model

$$f(dose) = a + (1-a)\times \boldsymbol{\Phi}\ (c + b\times \log(dose))$$

$$a\sim Uniform(0,1);\ b\sim Uniform(restriction,15);\ \ c\sim Uniform(-5,15)$$

Where “restriction” is a user defined value and can be 0, 0.5 or 1.

(8) Dichotomous Hill model

$$f(dose) = a\times g + \frac{a - a\times g}{1+e^{-c-b\times \log(dose)}}$$

$$a\sim Uniform(0,1);\ g\sim Uniform(0,1);\ b\sim Uniform(restriction,15);\ \ c\sim Uniform(-5,15)$$

Where “restriction” is a user defined value and can be 0, 0.5 or 1.

II.2.2 Continuous Data

For continuous data, there are eight models:

The background parameter “a” in all eight models has the same uniform distribution used as prior which is derived as follow:

The lower bound of the uniform distribution is always 0, and the upper bound is calculated differently for individual data and summary data. For individual data,

$$a_{upper} = \max(response) \times 2$$

i.e., doubling the largest response value in the input dataset.

For summary data,

$$a_{upper} = [\max(resp.mean) + 2 \times resp.sd_{mean.max}] \times 2$$

Where $\max(resp.mean)$ is the maximum mean response across all dose groups in the input dataset, and $resp.sd_{mean.max}$ is the response standard deviation in that dose group with the maximum mean response.

(1) Linear model

$$f(dose) = a + b \times dose$$

$$a \sim Uniform(0, a_{upper});\ b \sim Uniform(b_{lower}, b_{upper})$$

(2) Power model

$$f(dose) = a + b \times dose^{g}$$

$$a \sim Uniform(0, a_{upper});\ b \sim Uniform(b_{lower}, b_{upper});\ g \sim Uniform(restriction, 15)$$

Where "restriction" is a user defined value and can be 0, 0.5 or 1.

(3) Michaelis Menten model

$$f(dose) = a + \frac{b \times dose}{c + dose}$$

Current setting for model parameter priors are:

$$a \sim Uniform(0, a_{upper});\ b \sim Uniform(b_{lower}, b_{upper});\ c \sim Uniform(0, 15)$$

(4) Hill model

$$f(dose) = a + \frac{b \times dose^{g}}{c^{g} + dose^{g}}$$

Current setting for model parameter priors are:

$$a \sim Uniform(0, a_{upper});\ b \sim Uniform(b_{lower}, b_{upper});\ c \sim Uniform(0, 15);\ g \sim Uniform(restriction, 15)$$

Where "restriction" is a user defined value and can be 0, 0.5 or 1.

For the Linear, Power, Michaelis Menten and Hill model, the lower and upper bound of the parameter b (a slope-equivalent parameter) are determined by the dose-response trend and the overall slope in the input data.

For individual input data and increasing trend:

$$b_{lower} = 0$$

$$b_{upper} = \frac{Max(resp) - Min(resp)}{Dose_{Max\_resp} - Dose_{Min\_Resp}} \times 5$$

For individual input data and decreasing trend:

$$b_{lower} = \frac{Min(resp) - Max(resp)}{Dose_{Min\_resp} - Dose_{Max\_Resp}} \times 5$$

$$b_{upper} = 0$$

Where $Max(resp)$ and $Min(resp)$ are the maximum and minimum response value in the input dataset. And $Dose_{Max\_resp}$ and $Dose_{Min\_resp}$ are dose level corresponding to the maximum and minimum responses respectively.

For summary input data:

$$b_{slope} = \frac{Mean_{\max\ \_dose} + 2 \times SD_{\max\ \_dose} - Mean_{\min\ \_dose} - 2 \times SD_{\min\ \_dose}}{Max(dose) - Min(dose)}$$

Where $Mean_{\max\ \_dose}$ and $SD_{\max\ \_dose}$ are the mean and standard deviation of responses at the maximum dose level, and $Mean_{\min\ \_dose}$ and $SD_{\min\ \_dose}$ are mean and standard deviation of responses at the minimum dose level. $Max(dose)$ and $Min(dose)$ are the maximum and minimum dose level in the input dataset. Because dose levels are first normalized to the scale between 0 and 1, $Max(dose)$ is 1 and $Min(dose)$ is very likely 0. Then the prior distribution of the parameter "s" is $Uniform(0, 5 \times b_{slope})$ for increasing trend or $Uniform(5 \times b_{slope},\ 0)$ for decreasing trend.

(5) Exponential 2 model

$$f(dose) = a \times e^{b \times dose}$$

$a \sim Uniform(0, a_{upper})$; $b \sim Uniform(0, 50)$ for increasing trend or $Uniform(-50, 0)$ for decreasing trend.

(6) Exponential 3 model

$$f(dose) = a \times e^{b \times dose^g}$$

$a \sim Uniform(0, a_{upper})$;
$b \sim Uniform(0, 50)$ for increasing trend or $Uniform(-50, 0)$ for decreasing trend;
$g \sim Uniform(restriction, 15)$

Where "restriction" is a user defined value and can be 0, 0.5 or 1.

(7) Exponential 4 model

$$f(dose) = a \times [c - (c-1) \times e^{-b \times dose}]$$

$$a \sim Uniform(0, a_{upper});\ b \sim Uniform(0, 100);\ c \sim Uniform(0, 15)$$

(8) Exponential 5 model

$$f(dose) = a \times [c - (c-1) \times e^{-(b \times dose)^g}]$$

$$a \sim Uniform(0, a_{upper});\ b \sim Uniform(0, 100);\ c \sim Uniform(0, 15);\ g \sim Uniform(restriction, 15)$$

Where "restriction" is a user defined value and can be 0, 0.5 or 1.

II.3 Modeling Strategy

II.3.1 Dichotomous Data

For dichotomous data, three variables (i.e., dose level (d_i), total number of subjects in each dose group (n_i), and the number of subjects with effect in the corresponding dose group (y_i)) are needed to characterize the dose-response relationship based on the assumption that the number of subjects with effect at each dose group follow a binomial distribution $y_i \sim binomial\ (n_i, r_i)$, where the parameter $r_i = f(d_i|\boldsymbol{\theta})$ represents the probability of effect in the ith dose group and is determined by a dose-response function with a parameter vector of $\boldsymbol{\theta}$. Under a Bayesian framework, the regression analysis is essentially the process to estimate the posterior distribution of the model parameters $\boldsymbol{\theta}$, and thus,

$$p(\boldsymbol{\theta}|Data) \propto p(\boldsymbol{\theta}) \prod_{i=1}^{N} f(d_i|\boldsymbol{\theta})^{y_i} [1 - f(d_i|\boldsymbol{\theta})]^{n_i - y_i}$$

where $p(\boldsymbol{\theta}|Data)$ and $p(\boldsymbol{\theta})$ are the posterior and prior distribution of the model parameters respectively, N is the number of dose groups in the dataset, and $f(d_i|\boldsymbol{\theta})$ represents a parametric dose-response model. For individual dichotomous data, the individual data will be first converted to summary dichotomous data if there are more than one individual animals are tested at the same dose level.

II.3.2 Continuous Data

For continuous data (no matter they are individual or summary data), one fundamental assumption used in the BBMD system is that the continuous responses are lognormally distributed.

If individual response data are available, the dose level (d_i) and response (y_i) should be reported for each subject. So, the model fitting process estimates a posterior sample of model parameter $\boldsymbol{\theta}$ and the within dose group variance parameter γ^2, which can be expressed as,

$$p(\boldsymbol{\theta},\gamma|Data) \propto p(\boldsymbol{\theta},\gamma)\prod_{i=1}^{N}\frac{1}{\sqrt{2\gamma^2\pi}}\exp\left[-\frac{\{\log(y_i)-\log[f(d_i|\boldsymbol{\theta})]\}^2}{2\gamma^2}\right]$$

where $p(\boldsymbol{\theta},\gamma)$ is the prior distribution of the model parameters and variance parameter, and $f(d_i|\boldsymbol{\theta})$ represents a dose-response model for continuous data.

For summary continuous data, there are four reported variables: dose level (d_i), number of subjects in each group (n_i), the observed mean value of response in each group ($\bar{y}_i$), and the observed standard deviation of response in each group (s_i). Under the lognormality assumption, the commonly reported mean and standard deviation on regular scale are first converted to the corresponding quantities on a log scale by using $\bar{y}_i' = \log(\bar{y}_i) - 0.5\times\log\left[(s_i/\bar{y}_i)^2+1\right]$ and $s_i' = \sqrt{\log\left[(s_i/\bar{y}_i)^2+1\right]}$ (Crump 1995, Slob 2002). Modeling summary continuous data shares the same fundamental idea as the equation above but differentiated by input data, so the actual log-likelihood function that is used in the PyStan model for MCMC sampling is

$$\sum_{i=1}^{G} -\frac{n_i\log(\gamma^2)}{2} - \frac{n_i\times\{\bar{y}_i' - \log\left[f(d_i|\boldsymbol{\theta})\right]\}^2 + (n_i-1)\times s_i'^2}{2\gamma^2}$$

where G is the total number of dose groups.

II.4 Interpretation on Dose-Response Modeling Results

II.4.1 Textual Output of Parameters

The first box on the right panel of the "" tab (as shown in Figure 13) is directly obtained from PyStan's fit output, including some important statistics for model parameters and diagnostic indictors for the MCMC algorithms. For example, the "Rhat" can be used to judge if the MCMC chains have converged properly. If Rhat value is larger than 1.05, you may consider increase the length of MCMC chains to get better convergence. Detailed explanation on the Stan outputs can be found at: https://github.com/stan-dev/stan/releases/download/v2.9.0/stan-reference-2.9.0.pdf

II.4.2 Posterior Predictive P-value

A posterior predictive p-value is reported below the dynamic dose-response plot. This indicator can be used to judge if the fitting of this particular model is adequate. Practically, if the value is between 0.05 and 0.95, then the fitting is adequate. The calculation procedure is briefly described below:

(1) Use each bundle of parameters in the kept posterior sample to form a dose-response model and randomly generate case numbers, y^{rep}, at all dose levels in the original dataset
(2) Use posterior sample of model parameters to calculate a test statistic for both the original data set (d, n, y) and the replicated data set (d, n, y^{rep}). The test statistic used in this system is log-likelihood. For parameter values from the *l* th iteration, we have statistic $T(y, \theta^l)$ and $T(y^{rep}, \theta^l)$.
(3) For *l* = 1, … , L (the length of posterior sample), compare each pair of $T(y, \theta^l)$ and $T(y^{rep}, \theta^l)$, and count the number of $T(y, \theta^l) > T(y^{rep}, \theta^l)$, say *M*.
(4) The posterior predictive P-value is M/L

A detailed explanation on this procedure can be found in the Chapter of "Model checking and improvement" in *Bayesian Data Analysis* (Gelman et al).

II.4.3 Posterior Model Weight

A model weight is calculated and reported on this page. For each selected model, the value of $\widehat{m}$ is calculated as follow:

$$\widehat{m}_j = \exp\left[\hat{\boldsymbol{\ell}}_j - \frac{d_j}{2}\log(n)\right]$$

Where $\hat{\boldsymbol{\ell}}_j$ is a loglikelihood value using one set of posterior sample of parameter values, d_j is the number of parameters in the jth model, and n is the number of dose groups in the data set. So, the model weight can be calculated as (Wasserman, 2000)

$$\Pr(\boldsymbol{\mathcal{M}}_j | Data) = \frac{\widehat{m}_j}{\sum_{t=1}^{T} \widehat{m}_t}$$

This function assumes that equal model priors for all models selected, so the weight mainly indicates how well the model fit the data. To make the weight more reliable, we used 1000 sets of randomly selected posterior sample of model parameters to calculate the model weights. And the averaged model values are reported.

II.5 Model Averaged BMD Calculation

Before calculating model averaged BMD, a posterior sample of BMD from each individual model should be obtained. The process to get the posterior sample have been described in Section I.7. Here, we focus on the method to get model averaged BMD.

II.5.1 Posterior Model Weight for Model Averaged BMD Calculation

In this step, the prior model weight specified by users will be used in the posterior model weight calculation. The function is shown below. The $\widehat{m}$ for each model is calculated using the same procedure described in the previous section.

$$\Pr(\boldsymbol{\mathcal{M}}_j|Data) = \frac{\widehat{m}_j \Pr(\boldsymbol{\mathcal{M}}_j)}{\sum_{t=1}^{T} \widehat{m}_t \Pr(\boldsymbol{\mathcal{M}}_t)}$$

Based on the function above, we know that the posterior weight of a model will be 0 if the prior weight for the model is specified as 0.

II.5.2 Model Averaged BMD Calculation

For each model, we have posterior sample of BMD with the same length as the model parameters. Using default value, we should have:

$BMD_{1-1}, BMD_{1-2}, \dots, BMD_{1-15000}$ for model 1
$BMD_{2-1}, BMD_{2-2}, \dots, BMD_{2-15000}$ for model 2
…

Then, the posterior sample for model averaged BMD is calculated as:

$$BMD_{MA-1} = BMD_{1-1} \times w_1 + BMD_{2-1} \times w_2 + \cdots$$
$$BMD_{MA-2} = BMD_{1-2} \times w_1 + BMD_{2-2} \times w_2 + \cdots$$
…
$$BMD_{MA-15000} = BMD_{1-15000} \times w_1 + BMD_{2-15000} \times w_2 + \cdots$$

Therefore, we will have the same size of posterior sample for model averaged BMD. w_1, w_2, … are posterior model weight (prior model weight has been integrated) calculated from Section II.5.1.

References

Crump KS. 1995. Calculation of benchmark doses from continuous data. Risk Analysis, 15(1): 79-89.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B. (2003). Bayesian Data Analysis. Second Edition. Boca Raton, FL: Chapman and Hall/CRC Press.

Slob W. 2002. Dose-response modeling of continuous endpoints. Toxicological Science, 66(2): 298-312.

Wasserman, L. (2000). Bayesian model selection and model averaging. Journal of Mathematical Psychology, 44(1), 92-107.