A Web Based Dose-Response Modeling System – From Toxicological Data to Probabilistic RfD

Kan Shao [1], Andy Shapiro [2], and Weihsueh Chiu [3]

[1] School of Public Health, Indiana University, Bloomington, IN; [2] National Toxicology Program, RTP, NC; [3] College of Veterinary Medicine & Biomedical Sciences, Texas A&M University, College Station, TX

Contact: Kan Shao I Email: kshao@Indiana.edu I 812-856-2725

Introduction on Bayesian BMD (BBMD) Estimation

The benchmark dose (BMD) methodology has been recommended as a replacement to the traditional No (or Lowest) Observed Adverse Effect Level (NOAEL/LOAEL) method in quantitative human health risk assessment. Comparing to the NOAEL/LOAEL method, the BMD method is a more quantitative dose-response modeling approach in terms of the quantitative definition of adverse effect, the derivation of the lower bound of the BMD, etc.

As the mainstream of regulatory risk assessment is moving towards a probabilistic assessment framework, a standardized tool to perform both probabilistic BMD estimation and low-dose extrapolation is of urgent need. A core idea in probabilistic dose-response analysis is that the risk of having adverse effects at a specified dose level or the dose causing a certain level of risk should be probabilistically quantified and expressed in terms of distribution. The online BBMD system is designed to achieve those goals. This system is established primarily based on Bayesian statistical analysis featuring Markov Chain Monte Carlo (MCMC) algorithms in Stan Library for model fitting, parameter and quantity of interest (e.g., BMD level) estimation. After distributional BMD is estimated, Monte Carlo (MC) simulation is employed for low-dose extrapolation to derive probabilistic RfD.

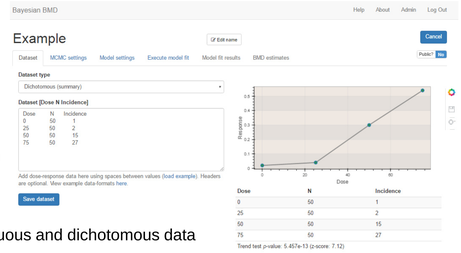
Available at https://benchmarkdose.org

Features of the BBMD System

The updated system includes the following steps (organized in tabs): (1) Dataset; (2) MCMC setting; (3) Model settings; (4) Execute model fit; (5) Model fit results; (6) BMD estimates and (7) Probabilistic RfD calculation

Dataset

- Four types of dose-response data can be analyzed: dichotomous summary; dichotomous individual; continuous summary; and continuous individual
- Input box accepts data: data can be manually input or copied and pasted from other sources (ex: Microsoft Excel)
- The system adopts data formats commonly reported in literature for continuous and dichotomous data



- Once the dataset is successfully saved, dataset can be retrieved for future use. The data will be visualized (not fitted) in a plot and table. Additionally, a trend test will be performed for dichotomous data to report if the data have a clear trend for dose-response modeling

MCMC settings

- The length of a MCMC with a default of 30,000 (10,000 – 50,000)
- The number of chains with a default value of 1 (1 – 3 chains)
- The warmup percentage for discarding sample as burn-in with a default of 50% (10 – 90%)
- Random seed randomly generated or can be specified for reproduction

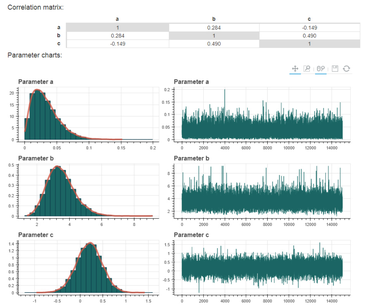
Model settings

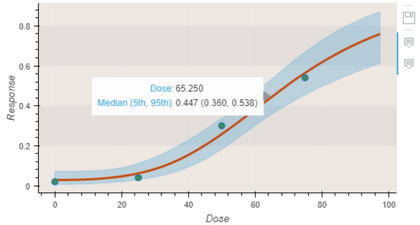
- Eight dichotomous and eight continuous dose-response models currently available for selection:
  - **Dichotomous:** Logistic, LogLogistic, Probit, LogProbit, Quantal Linear, Multistage (2nd), Weibull, Dichotomous Hill
  - **Continuous:** Exponential (2, 3, 4, 5), Hill, Power, Michaelis Menten, Linear

Execution and model fit results

- A step to check all required inputs are successfully validated, then press "Execute" to fit the models
- The fit-results including both textual and graphical outputs are available for each of the selected models
- Textual output summarizes important statistics regarding model parameters with indicators representing the quality of chain convergence
- A covariance matrix is calculated and presented
- A distribution plot and posterior sample trace plot for each model parameter can be shown or hidden

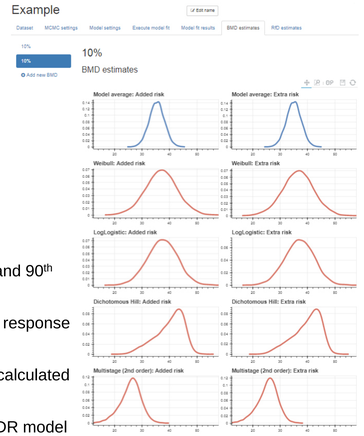
```
Inference for Stan model: anon_model_20bb8bf569f4822af7976eed54022d1e.
1 chains, each with iter=30000; warmup=15000; thin=1;
post-warmup draws per chain=15000, total post-warmup draws=15000.

        mean se_mean     sd   2.5%    25%    50%    75%  97.5%  n_eff   Rhat
a       0.03  2.0e-4   0.02 4.5e-3   0.02   0.03   0.04   0.08  10395    1.0
b        3.5    0.01   0.88   2.03   2.89   3.41   4.01   5.42   7261    1.0
c        0.2  3.1e-3   0.28  -0.36   0.01    0.2   0.38   0.75   8046    1.0
lp__ -83.09    0.02    1.3 -86.48 -83.67 -82.75 -82.14 -81.62   5925    1.0

Samples were drawn using NUTS at Tue Dec  5 12:35:05 2017.
For each parameter, n_eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor on split chains (at
convergence, Rhat=1).
```





Posterior predictive *p*-value for model fit: 0.598

Model weight: 33.4%



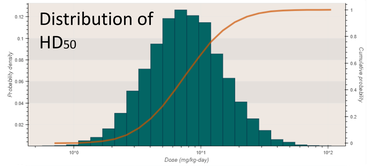
- Interactive DR curve is plotted, with median DR curve and 90th percentile interval are shaded
- Moving the mouse on the curve will display the median response with 5th and 95th percentile at each dose level
- Posterior predictive p-value (PPP) for a test statistic is calculated to indicate the adequateness of the model fitting
- Model posterior weight is calculated for each selected DR model for model comparison

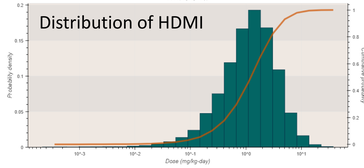
BMD estimation

- For dichotomous data, BMD estimates are calculated based on BMR definitions of extra risk & added risk
- For continuous data, various definitions of BMR are used for BMD estimation, including absolute and relative change in central tendency (i.e., median), and hybrid approach based on tails of a distribution
- Both model averaged BMD and individual model BMD are calculated
- User can specify prior model weight for selected models for model averaged BMD calculation (or remove models by specifying a prior-weight of 0)
- Distributional BMD estimates are shown in textual and graphical outputs

Probabilistic RfD Calculation

- The posterior sample of BMD estimates is used as the POD for probabilistic low-dose extrapolation
- Various sources of uncertainty and variability are considered, including uncertainty in animal-to-human TK/TD difference, uncertainty in extrapolation of exposure duration, human variability, etc.
- These uncertainties and variabilities are assumed to be lognormally distributed; and samples are generated using Monte Carlo simulation
- HD_{50}, estimated human dose at which 50% of the population has effects greater than or equal to the target magnitude (M) of effect is estimated
- HDMI, estimated human dose where the population has I % incidence (including inter-individual human variability) with effects greater than or equal to target magnitude (M)
- The contribution of each uncertainty/variability factor to overall uncertainty is calculated
- Target Incidence Level I* vs Human Dose plot is produced to show the confident intervals of human dose that can achieve targeted protection goals



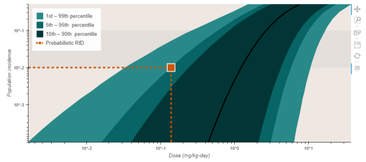


Uncertainty contributions

Ratio of 95th percentile to 5th percentile in HDMI: 41.5

Uncertainty	Contribution
Variability within point of departure	0.672%
Allometric scaling	2.04%
Test species to human toxicokinetics/toxicodynamics	57.5%
Duration of exposure	-
Other uncertainty #1	-
Other uncertainty #2	-
Human variability	39.8%

Target incidence level I* vs. Human dose



Future Work

- Allow user to specify prior distribution for parameters of different dose-response models
- Build a comprehensive database that can be used to elicit more toxicologically plausible prior distribution for model parameters to enhance the reliability of and efficiency of BMD estimation