A Next Generation Benchmark Dose Computation System

Kan Shao [1] and Andy Shapiro [2]

[1] Assistant Professor, Department of Environmental Health, School of Public Health, Indiana University – Bloomington; [2] Independent Consultant, Durham, NC

School of Public Health Bloomington
INDIANA UNIVERSITY

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

The Need of Bayesian BMD Estimation

The benchmark dose (BMD) methodology has mostly replaced 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, and etc.

There are two major BMD estimation software currently available: (1) US EPA's Benchmark Dose Software; and (2) the PROAST software published by the Netherlands' National Institute for Public Health and the Environment. Both software played an important role in facilitating and promoting the BMD methodology. The statistical methodology used in the current software is mainly based on the traditional frequentist statistics, i.e., the maximum likelihood estimation (MLE) algorithms are implemented to get point estimates for parameters of dose-response models and BMD. Although a lower bound of the BMD estimate is calculated, it is not a comprehensive description for the distribution of the BMD estimate.

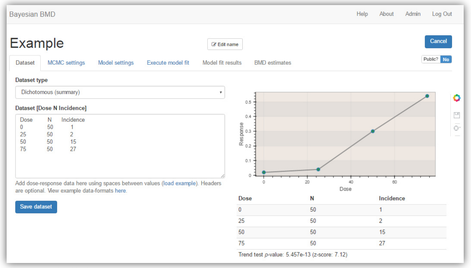
As the mainstream of regulatory risk assessment is moving towards a probabilistic assessment framework, a standardized tool to perform both probabilistic BMD estimation and low-dose extrapolation is of urgent need. A core idea in probabilistic dose-response analysis is that the risk of having adverse 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 BMD estimation system being developed is exactly designed to achieve those goals. This system is established primarily based on Bayesian statistical analysis featuring Markov Chain Monte Carlo (MCMC) algorithms for dose-response model fitting, parameter and quantity of interest (e.g., BMD level) estimation.

Demo: http://benchmarkdose.com

Features of the Online BBMD System

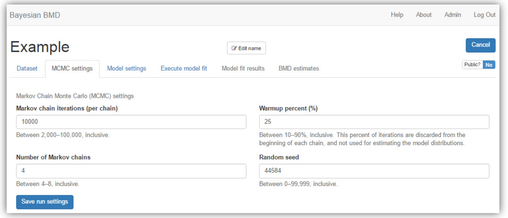
This new system includes following sections/steps: (1) Dataset; (2) MCMC setting; (3) Model settings; (4) Execute model fit; (5) Model fit results; and (6) BMD estimates

Dataset



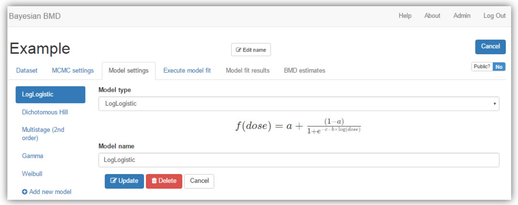
- User selects a dose-response data type for analysis: dichotomous summary; dichotomous individual; continuous summary; and continuous individual
- Input box accepts data: data can be manually input or copied and pasted from other sources (ex: Microsoft Excel)
- The system adopts data formats commonly reported in literature for continuous and dichotomous data
- Once user presses the "Save dataset" button, the dataset will be saved 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 10,000 (2,000 – 100,000)
- The number of chains with a default value of 4 (4 – 8 chains)
- The warmup percentage for discarding sample as burn-in with a default of 25% (10 – 90%)
- Random seed randomly generated or can be specified for reproduction
- Press "Save run settings", then all MCMC settings will be saved

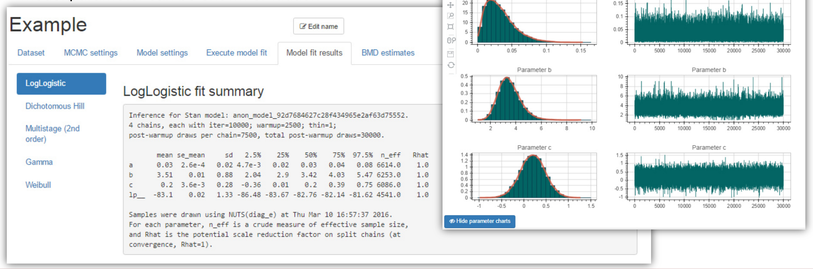
Model settings



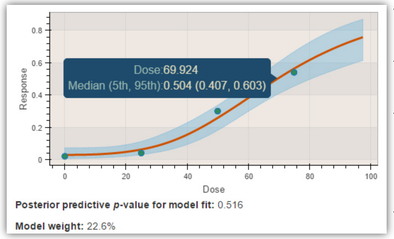
- Nine dichotomous and eight continuous dose-response models currently available for selection:
  - **Dichotomous:** Gamma, Logistic, LogLogistic, Probit, LogProbit, Quantal Linear, Multistage (2nd), Weibull, Dichotomous Hill
  - **Continuous:** Exponential (2, 3, 4, 5), Hill, Power, Michaelis Menten,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

Execution and model fit results

- A step to check all require inputs are successfully saved, then press "Execute" to fit the models
- Fitting results are available for each of the selected models. The fit-results are shown on the right panel, including both textual and graphical outputs
- The first part of 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



BMD estimation



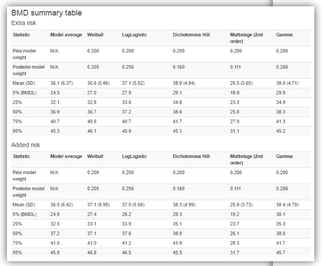
- Interactive DR curve is plotted, with median DR curve and 90th percentile interval are shaded
- Moving the mouse on the curve will display the median response with 5th and 95th percentile at each dose level
- Posterior predictive p-value (PPP) for a test statistic is calculated to indicate the adequateness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

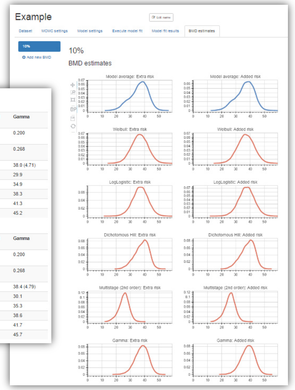
$$\Pr[T(y,\theta^l) > T(y^{rep,l},\theta^l)]$$

- Model posterior weight is calculated for each selected DR model for model comparison using the equations below:

$$\Pr(M_j|Data) = \frac{m_j}{\sum_{r=1}^{R} m_r} \qquad m_j = \exp(\hat{\ell}_j - \frac{d_j}{2}logn)$$

- Various definitions can be used to calculate BMD for dichotomous and continuous data
- Calculate model averaged BMD and individual model BMD
- User can specify prior model weight for selected models for model averaged BMD calculation (or remove models by specifying a prior-weight of 0)
- Textual and graphical outputs of BMD estimates are shown





Technical implementation

- Web-application created using the Python Django web-framework & the React.js frontend for user-interface
- Documented RESTful API can be used by other applications for batch modeling applications
- Permissions schema allows for runs to be shared publicly, should a user wish to make results public

Improvements in the Future

Additions in the near future

- Allow user to put restrictions on power parameters in multiple dose-response models
- Allow user to specify prior distribution for parameters of dose-response models
- Calculate model averaged risk at any user-specified dose level
- Generate downloadable reports (Microsoft Word) of summary results

Extended implementations

- Build a prior distribution database (estimated from key studies of various chemicals) for model parameters with the potential to utilize new toxicological data