Interpreting the results of EPA dose-response models

Jeff Gift, PhD
J. Allen Davis, MSPH
National Center for Environmental Assessment, U.S. EPA
October 24, 2017
Peer Review of Draft NTP Approach to Genomic Dose-Response Modeling



Disclaimer

The views expressed in this presentation are those of the author(s) and do not necessarily reflect the views or policies of the US EPA.

Outline

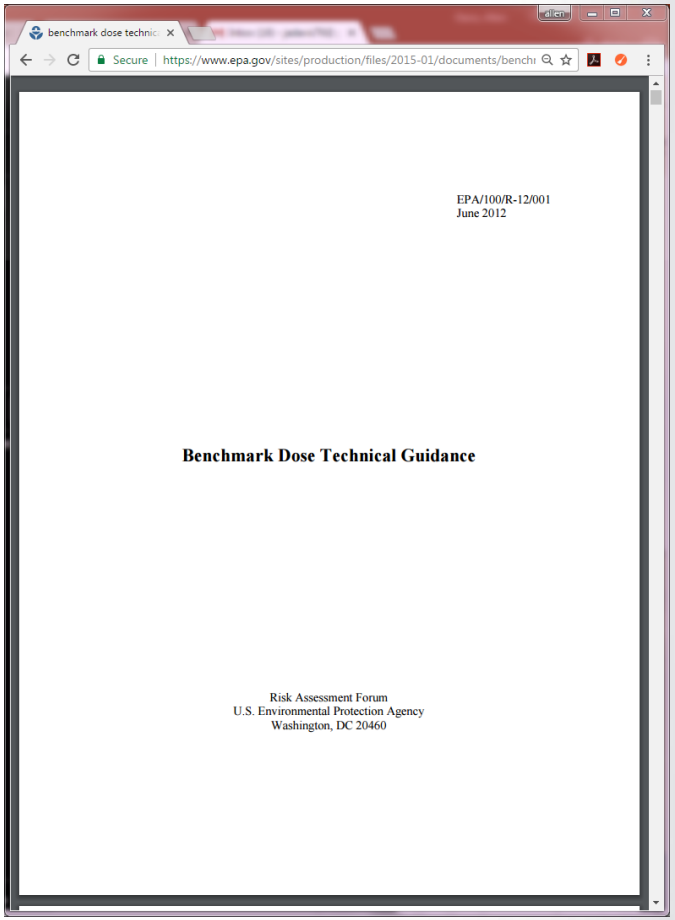
- **Existing BMD modeling approaches for continuous data**
- **Future directions for BMDS**
- **NTP's BMDExpress**

EPA's BMD Technical Guidance

- ***Final*** **version of the EPA's Benchmark Dose (BMD) Technical Guidance document was published in 2012: https://www.epa.gov/risk/benchmark-dose-technical-guidance**

- **Other guidance documents relevant to BMD modeling available at: http://epa.gov/iris/backgrd.html**

- **EPA's Statistical Working Group periodically updates recommended model practices**



BMD Analysis Key Steps

1. **Benchmark Response (BMR):** Choose BMR(s) to evaluate.

2. **Model Selection:** Run appropriate models and parameter options.

3. **Model Fit:** Determine which models adequately fit the data.

4. **BMDLs:** Are BMDLs for models that fit *sufficiently close* (< 3-fold)*?*

5. **Akaike Information Criterion (AIC):** If so, use model with lowest AIC; if not use model with lowest BMDL

BMR: Continuous Data

- **Preferred approach: Use a benchmark response (BMR) that corresponds to a level of change representing a minimal biologically significant response (i.e., 10% decrease in body wt.)**
- **In the absence of a biological consideration, a BMR of a change in the mean equal to one control standard deviation (1.0 SD) from the control mean is recommended.**
- **In some cases, use of different BMRs is supported (e.g., 0.5 SD for changes in critical organ systems)**

Continuous Model Selection

Biological Interpretation	**Example: Hill or Exponential models can be used for receptor-mediated responses**
Policy Decision	**U.S. EPA's OPP group uses the Exponential models for modeling acetylcholinesterase inhibition data**
Otherwise	**In the absence of biological or policy-driven considerations, criteria for final model selection are usually based on whether various models mathematically describe the data**

Continuous Model Forms

Model Name	Functional Form	# of Parameters	Model Fits
Polynomial[a]	$\beta_0 + \beta_1 X + \beta_2 X^2 + \ldots + \beta_n X^n$	1 + n	All purpose, can fit non-symmetrical S-shaped datasets with plateaus
Power	$\gamma + \beta X^{\Phi}$	3	L-shaped
Hill	$\gamma + \frac{(v \times X^n)}{(k^n + X^n)}$	4	Symmetrical, sigmoidal, S-shape with plateau
Exponential[b] Model 2 Model 3 Model 4 Model 5	$a \times \exp\{\pm 1 \times b \times X\}$ $a \times \exp\{\pm 1 \times (b \times X)^d\}$ $a \times [c - (c-1) \times \exp\{\pm 1 \times b \times X\}]$ $a \times [c - (c-1) \times \exp\{\pm 1 \times (b \times X)^d\}]$	 2 3 3 4	All purpose (Models 2 & 3) Symmetrical and asymmetrical S-shape with plateau (Models 4 & 5)

[a] The stand-alone Linear model in BMDS is equal to a first-order polynomial model

[b] Nested family of 4 related models described by Slob (2002) and included in the PROAST software of RIVM

Restricting Model Parameters

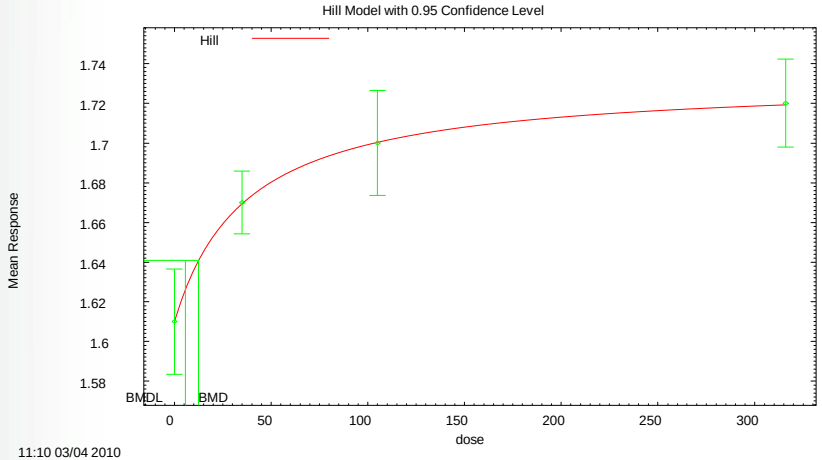
- **Model parameters (i.e., slope, background response, etc.) are generally bounded to restrict the dose-response curve to "reasonable" shapes**
- **These restrictions can impact statistical calculations such as the goodness-of-fit p-value and Akiake AIC**
- **The use of model restrictions is a topic of ongoing discussion in EPA's Statistical Working Group**

Does the Model Fit the Data?

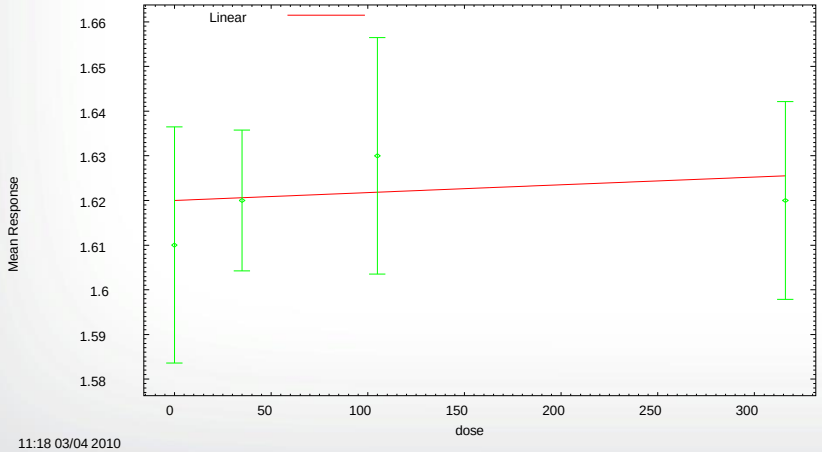
- **Tests of interest (response/variance modeling)**
- **Global measurement: p>0.1; p>0.05 for preselected models**
- **Local measurement: Scaled residuals (absolute value < 2)**
- **Visual inspection of model fitting.**
- **BMD/BMDL: caution if > 5; serious concern if > 20 (Wizard)**
- **Use of AIC or BMDL range to choose between adequately fitting models**

Tests of Interest – Differences in Responses and/or Variances

- **Test 1 – Do responses and/or variances differ among dose levels?**



"The p-value for Test 1 is less than .05. There appears to be a difference between response and/or variances among the dose levels. It seems appropriate to model the data"



"The p-value for Test 1 is greater than .05. There may not be a difference between responses and/or variances among the dose levels. Modeling the data with a dose-response curve may not be appropriate"

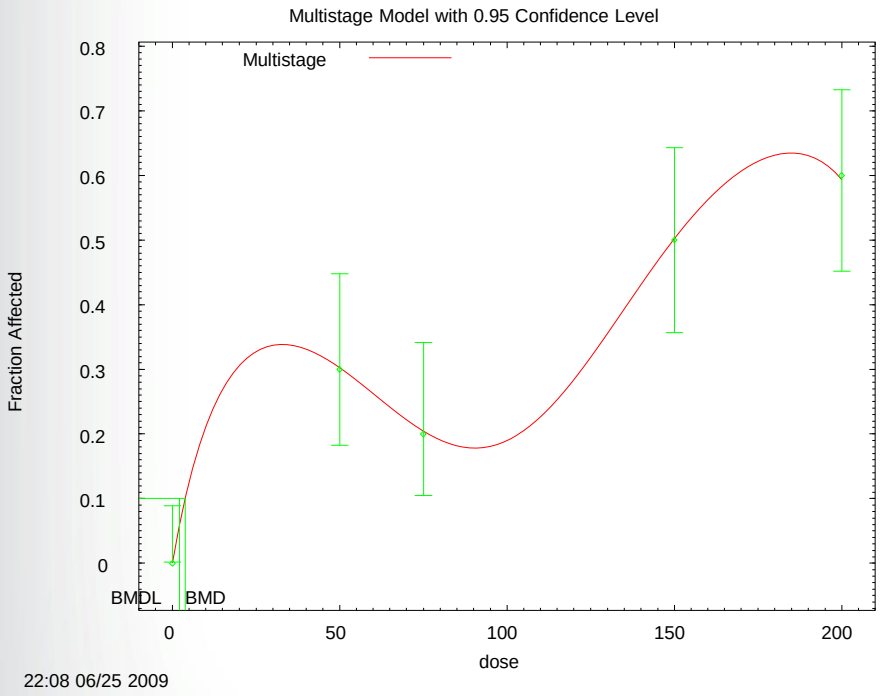
Tests of Interest – Variance

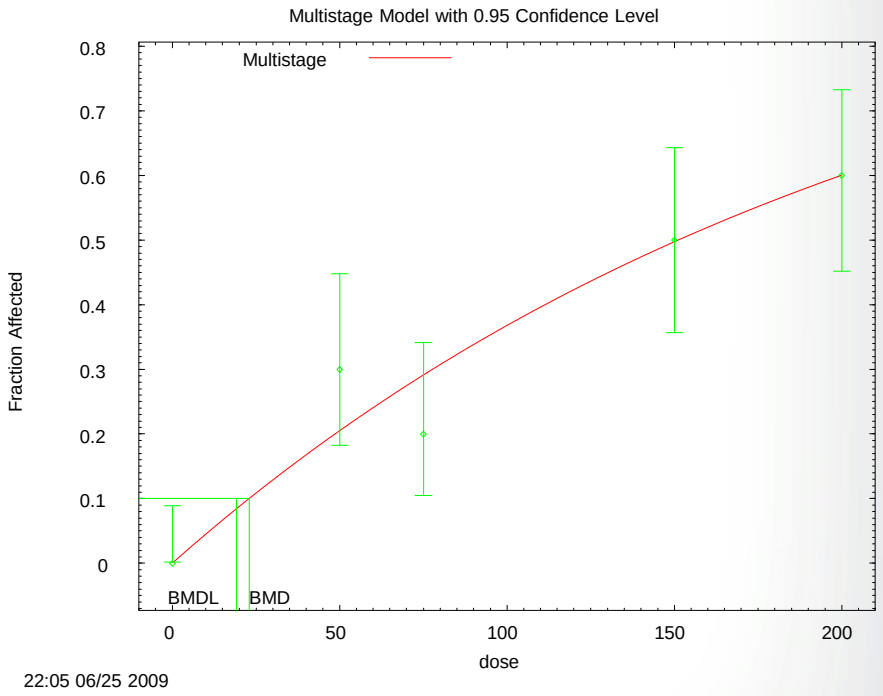
- **Distribution of continuous measures is assumed to be normal, with either a constant (homogenous) variance or a variance that changes as a power function of the mean**
  - $Var(i) = \alpha[mean(i)]^{\rho}$
  - ρ(rho) = 0, constant variance
  - ρ(rho) $\neq$ 0, modeled variance
- **Test 2 – Are variances homogenous?**
- **Test 3 – Are variances adequately modeled?**
- **Recommendation is to assume constant variance unless data clearly indicate otherwise**

Goodness-of-Fit

- **Global - BMDS provides *p*-value to measure *global* goodness-of-fit**
  - Measures how model-predicted dose-group probability of responses differ from the actual responses
  - Small values indicate poor fit
  - Cut-off value is $p = 0.10$ (0.05 when selecting a model *a priori*)
- **Local - Scaled Residuals measure the *local* fit of the model at each point; 0 = exact fit**
  - Continuous data: $\frac{Obs\ Mean - Est\ Mean}{\frac{Est\ SD}{\sqrt{n}}}$
  - Absolute values near the BMR should be lowest
  - Question scaled residuals with absolute value > 2

Visual Inspection of Fit





Options if Goodness-of-Fit is Poor

- **What to do when goodness-of-fit is poor?**
  - Use a different model
  - Consider dropping high dose group(s)
  - Use alternative dose metric (e.g., internal doses from PBPK model)
  - Log-transform doses if appropriate

Are BMDL Estimates "Sufficiently Close"?

- **What is "sufficiently close" can vary based on needs of the assessment, but generally should not be more than 3-fold.**
  - If BMDLs are not sufficiently close, ***EPA recommends picking the model with the lowest BMDL***
  - If BMDLs are sufficiently close, ***EPA recommends selecting the model with the lowest AIC***
  - If multiple models have the same AIC, ***EPA recommends combining BMDLs***

Future of BMDS

- **New alternative approaches to the "best model" method:**
  - Model Averaging – Bayesian or frequentist
  - Use of single, hyper-flexible model (parametric or non-parametric)
  - Preliminary tests for continuous model averaging methods indicate coverage is better than "best model" approach

Source: Wheeler, Gift, and Davis (2017). In prep.

NTP Proposed Approaches to Genomic Dose-Response Modeling

NTP Proposed Approach to Genomic Dose-Response Modeling

Four Step Process

1. Filtering the measured features (genes/probe sets)
2. Fitting filtered features to dose-response models
    a. Selecting and running models
    b. Evaluating model fit
3. Parsing the features into predefined gene sets (e.g., Gene Ontology Biological Processes); and
4. Determining potency for each of the adequately populated gene sets by deriving the mean and median potency of the genes in each set.

Draft Document: https://ntp.niehs.nih.gov/ntp/about_ntp/ntpexpanel/2017/october/approach20171025_508.pdf

NTP Proposed Approach to Genomic Dose-Response Modeling

Step 1: Filtering the measured features

NTP Proposal	Analogous EPA Approach
One-way ANOVA, $p < 0.05$ and fold-change (empirically determined for each technology)	Trend test generally preferred; may exclude more datasets than ANOVA test
Trend test not recommended because "biologically meaningful, non-monotonic responses can be observed that would be removed by a simple trend test."	Non-monotonic responses considered unlikely for typical BMD endpoints

NTP Proposed Approach to Genomic Dose-Response Modeling

Step 2a: Selecting and Running Models

NTP Proposal	Analogous EPA Approach
Hill, power, Linear, Polynomial 2° and 3°, Exponential 2,3, 4 and 5	Same
Response data are log2 adjusted and presumed to have normal distribution	Normal or lognormal response distribution allowed for Exponential models (for all models in next BMDS version)
Model runs assume constant variance (future version to test and compare fit of constant and non-constant variance models)	Compare fit of constant and non-constant variance models; Choice of variance model can impact BMD/BMDL estimation

NTP Proposed Approach to Genomic Dose-Response Modeling

Step 2a: Selecting and Running Models (continued)

NTP Proposal	Analogous EPA Approach
BMR = 1.349 X SD at 0 dose; ~10 % Δ in transcript abundance vs control	BMR set to biologically relevant change or 1.0 SD by default.
SD estimated from entire fitted curve, not just the control data	Same; important to consider model estimate of control SD
For most models, adverse direction determined by BMDS model	Adverse direction determined by BMDS or set by user
Adverse direction of Polynomial 2° and 3° models determined by which direction results in lowest BMD	BMDS uses linear trend test to auto-detect adverse direction. Care is warranted if data are non-linear or non-monotonic.

NTP Proposed Approach to Genomic Dose-Response Modeling

Step 2a: Evaluating Model Fit

NTP Proposal	Analogous EPA Approach
Exclude feature if no model has a complete set of convergent BMD, BMDU and BMDL values	Not a likely scenario for the limited number of datasets analyzed in an IRIS assessment
Use nested likelihood ratio test to choose between adequate fitting (BMDS p>0.05) linear and polynomial models • Choose more complex model if fit is significantly improved ($p < 0.05$) • Otherwise, choose simpler model	EPA uses AIC to choose between models and has not proposed a nested selection approach such as this for continuous polynomial models
Compare AIC for chosen Polynomial model with AIC for adequate fitting power, Hill, and exponential models	Consistent with EPA BMD "best model" approach minus "BMDL range" consideration

NTP Proposed Approach to Genomic Dose-Response Modeling

Step 2b: If Hill model is best fitting model

NTP Proposal	Analogous EPA Approach
Accept if Hill model "k" parameter is >1/3 the lowest positive dose	Accept if BMD/BMDL ratio is <20; serious concern if higher
If Hill model "k" parameter is <1/3 of the lowest positive dose select next best model with fit p-value >0.05	Use caution if BMD/BMDL ratio is >5, "serious concern" (probably reject) if above 20
If no model has a p >0.05, assign BMD equal to lowest BMD from the probe set with a acceptable Hill model ("k" parameter>1/3 lowest positive dose)	If a model gives a BMD/BMDL ratio of > 20 and no other model has a p >0.1, EPA might relax fit criteria (e.g., to p >0.05) or use another dataset.

NTP Proposed Approach to Genomic Dose-Response Modeling

Step 3: Filtering of features before gene set analysis

NTP Proposal	Analogous EPA Approach
Remove features with fit p-value<0.0001	Remove features with fit p-value<0.05 (pre-selected model) or <0.1 (no preselected model)
Remove features with BMDU/BMDL > 40	Remove features with BMD/BMDL > 20

Conclusions

- **BMD modeling for traditional continuous endpoints using "best method" approach used for 20+ years; methods are well-defined**
- **Alternative methods are being researched to address model uncertainty (e.g., model averaging) and provide more accurate modeling results**
- **BMDExpress leverages BMDS model executables to extend methods to alternative endpoints (i.e., gene expression)**
- **BMDExpress modeling and model selection criteria are generally consistent with EPA methods in areas of overlapping purpose.**
- **BMDExpress with well positioned to adapt to updates to BMD modeling approaches (i.e., adoption of model averaging)**