



Dose-Response Modeling for Risk Assessment – BMDS 3.2

Allen Davis, MSPH

Jeff Gift, Ph.D.

Center for Public Health and Environmental Assessment, U.S. EPA

Matt Wheeler, Ph.D.

National Institute for Occupational Safety and Health





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 or NIOSH.



Contributors – Software and Training Development

- **U.S. EPA Center for Public Health and Environmental Assessment (CPHEA)**
 - Jeffrey Gift, Ph.D. (gift.jeff@epa.gov; 919-541-4828)
 - J.Allen Davis, MSPH (davis.allen@epa.gov; 205-422-0655)
 - Todd Blessinger, Ph.D.
 - Matthew Wheeler, Ph.D. (on detail from NIOSH)
 - Jay Zhao, Ph.D.
- **General Dynamics Information Technology (GDIT)**
 - Cody Simmons
 - Louis Olszyk
 - Michael Brown
- **Bruce Allen Consulting**
 - Bruce Allen, M.S.



U.S. EPA Benchmark Dose Technical Guidance

- **EPA's Benchmark Dose Technical Guidance document:**
<http://www.epa.gov/raf/publications/benchmarkdose.htm>
- **Other guidance documents relevant to BMD modeling available at:**
<http://epa.gov/iris/backgrd.html>
- **BMDS User Guide, technical memos, glossary of terms and more at**
<https://www.epa.gov/bmds>



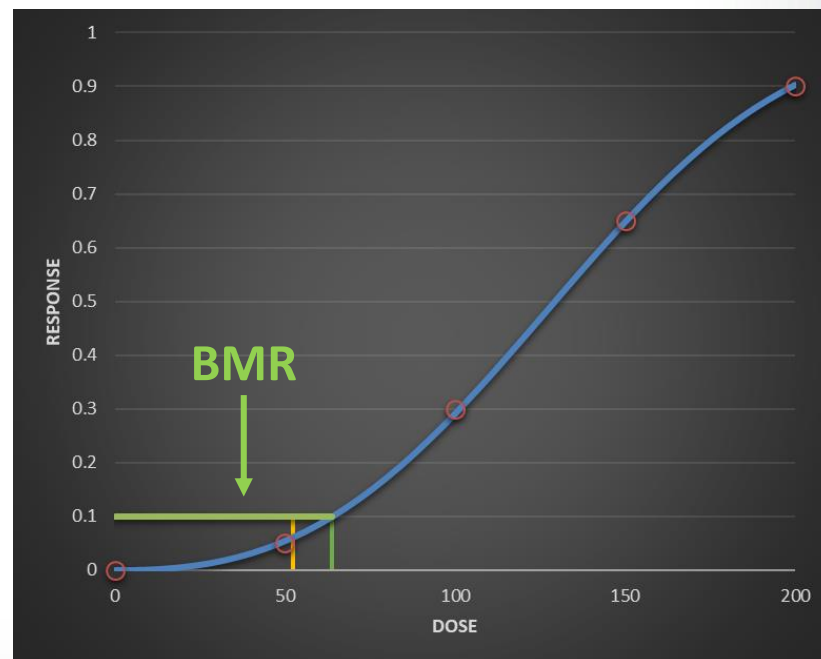
A Brief History of the BMD Method

1983	EPA workshop on epigenetic carcinogenesis
1984	“Benchmark dose” coined by Kenneth Crump in “A new method for determining allowable daily intakes”
1985–1994	Several EPA BMD-related publications and workshops
1995	EPA Risk Assessment Forum discusses use of BMD in risk assessment
1995	First IRIS BMD-based RfD (Methylmercury)
2000	EPA benchmark dose draft technical guidance released
2000	EPA benchmark dose software (BMDS) released
2000–2011	Multiple versions of BMDS released
2012	EPA benchmark dose final technical guidance released
2018/19	Incorporation of Bayesian model averaging into BMDS



Benchmark Dose – Key Terminology

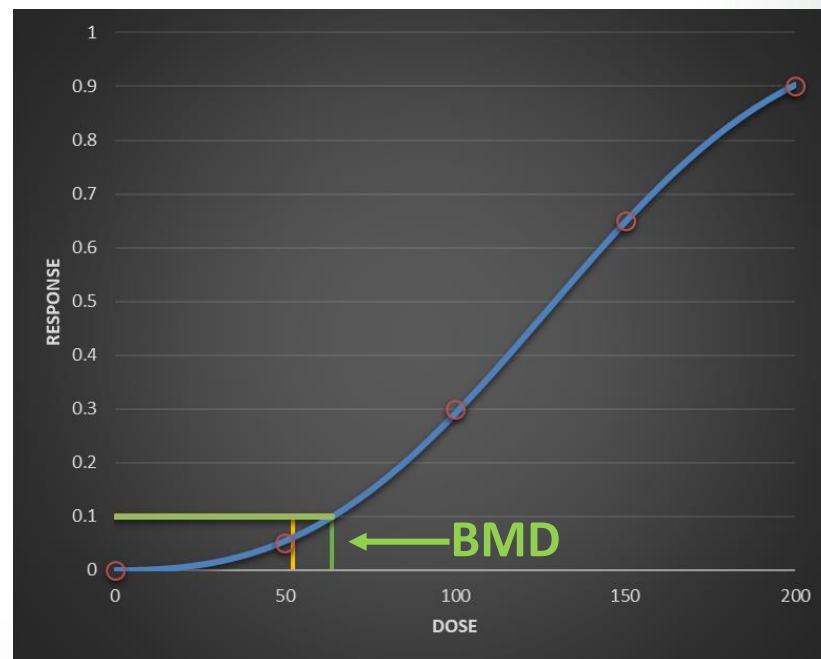
- **Benchmark Response (BMR)** - a change in response for an effect relative to background response rate of this effect
 - Basis for deriving BMDs
 - User defined
- **Examples include:**
 - 1 standard deviation increase in body weight (continuous response)
 - 10% increase in hepatocellular hyperplasia (dichotomous response)





Benchmark Dose – Key Terminology

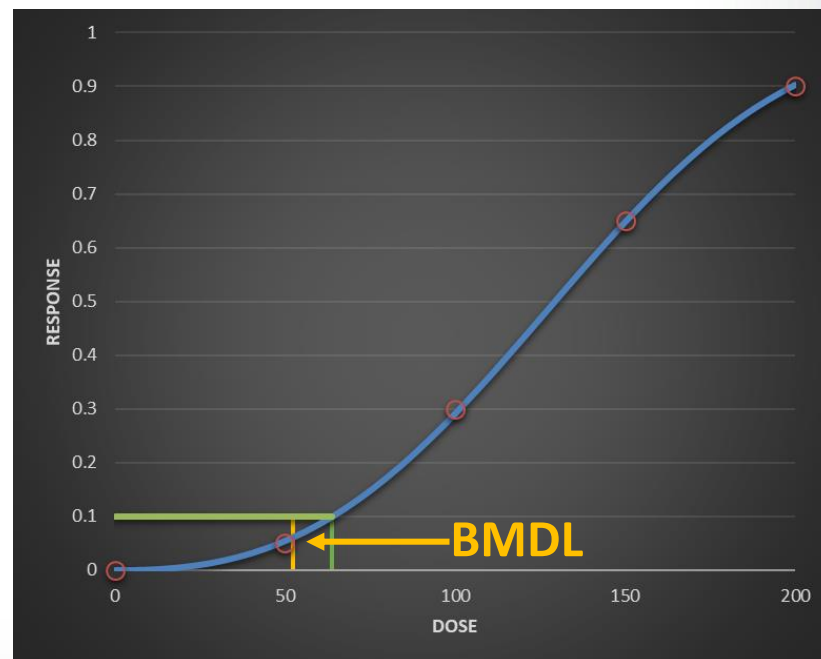
- **Benchmark dose or concentration (BMD or BMC)** - the maximum likelihood estimate of the dose associated with a specified benchmark response level
 - BMD – oral exposure
 - BMC – inhalation exposure
- **However, the term benchmark dose modeling is frequently used to describe the modeling process for both oral and inhalation exposures.**





Benchmark Dose – Key Terminology

- **Benchmark dose or concentration lower-confidence limit (BMDL or BMCL)** – the lower limit of a one-sided confidence interval on the BMD (typically 95%)
 - BMDL – oral exposure
 - BMCL – inhalation exposure
- **Accounts for elements of experimental uncertainty, including:**
 - Sample size
 - High background response
 - Response variability
- **Preferred POD**



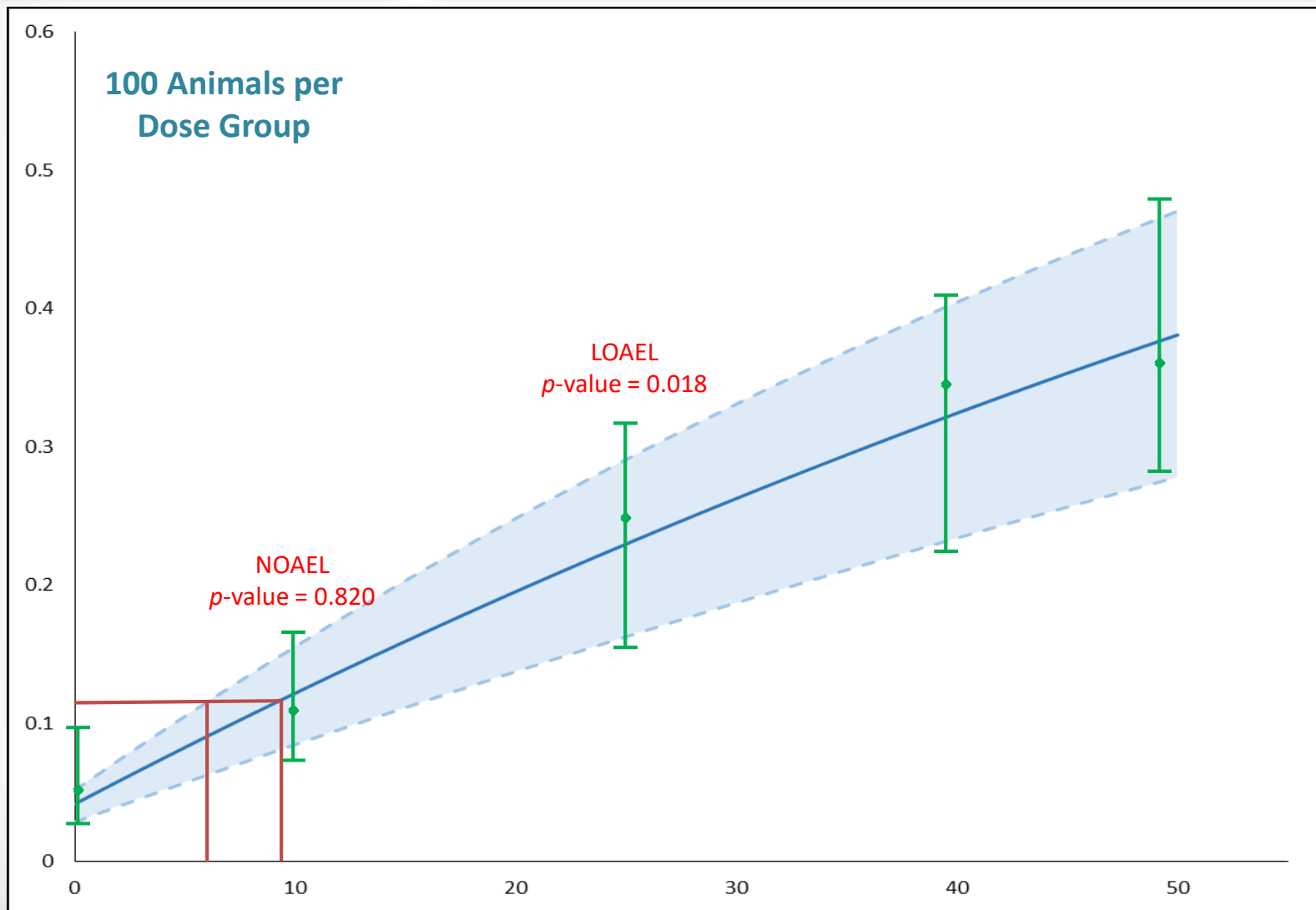


Advantages of BMD Approach

Subject	BMD Approach
Dose selection	BMD and BMDL not constrained to be a dose used in study
Sample size	Appropriately considers sample size: as sample size decreases, uncertainty in true response rate increases (i.e., $\downarrow N = \downarrow \text{BMDL}$)
Cross-study comparison	Observed response levels at a selected BMR are comparable across studies (recommended to use BMD as point of comparison)
Variability and uncertainty in experimental results	Characteristics that influence variability or uncertainty in results (dose selection, dose spacing, sample size) are taken into consideration
Dose-response information	Full shape of the dose-response curve is considered
NOAEL not identified in study	A BMD and BMDL can be calculated even when a NOAEL is missing from the study

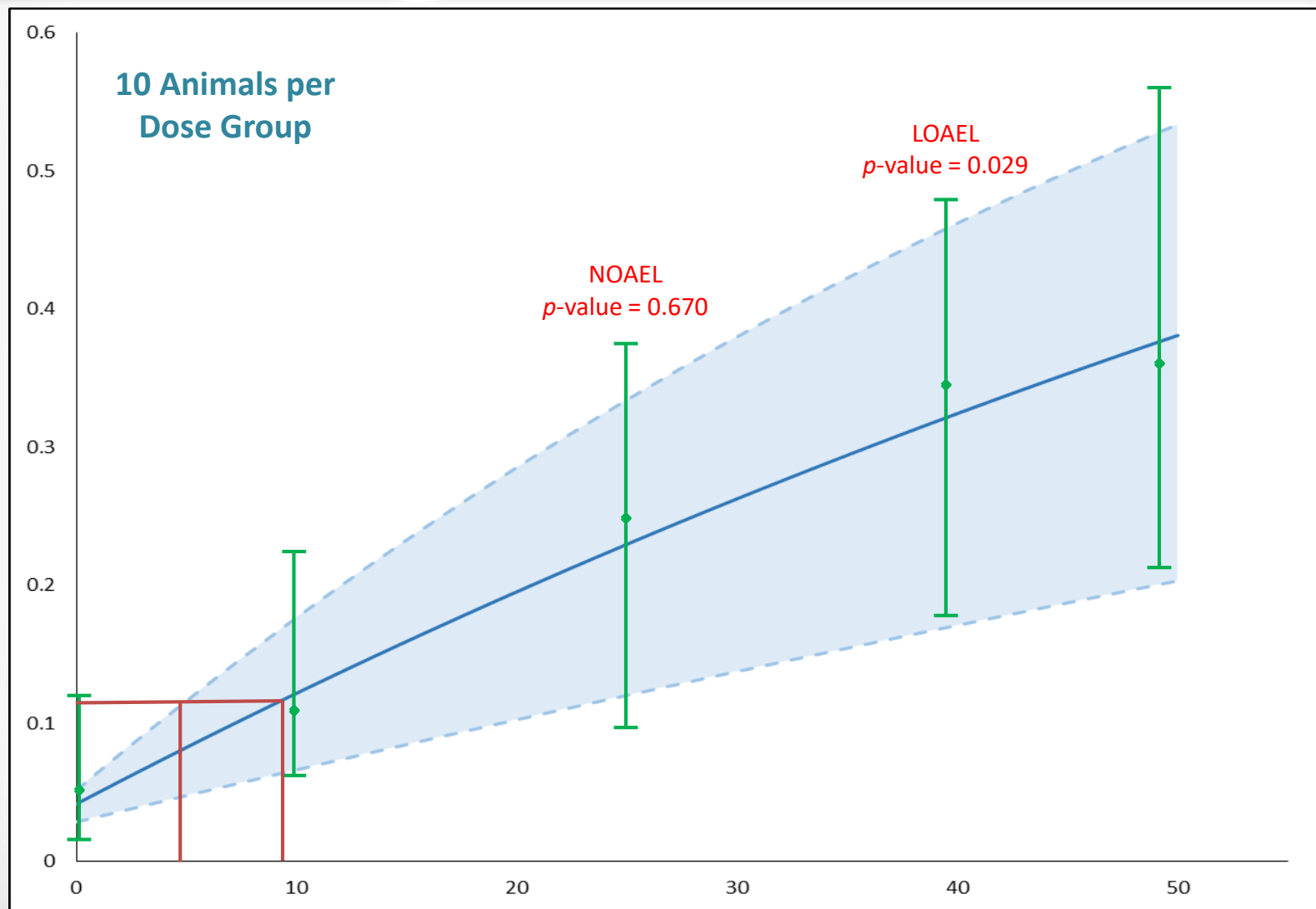


Study Size Effects on BMD/BMDL Calculation





Study Size Effects on BMD/BMDL Calculation





Challenges in the Use of the BMD Method

- **Requires knowledge on how to use software and interpret results**
- **In some cases, more data are required to model benchmark dose than to derive a LOAEL/NOAEL**
 - Continuous data require a measure of variability (SD or SE) for each dose group's mean response
 - Individual animal-level data are required for some models
 - Results highly dependent on the quality of the data
- **Sometimes the data cannot be adequately fit by the available models in BMDS**
- **Currently “best” model selection can add complexity and subjectivity to dose-response analyses**



U.S. EPA's Benchmark Dose Software

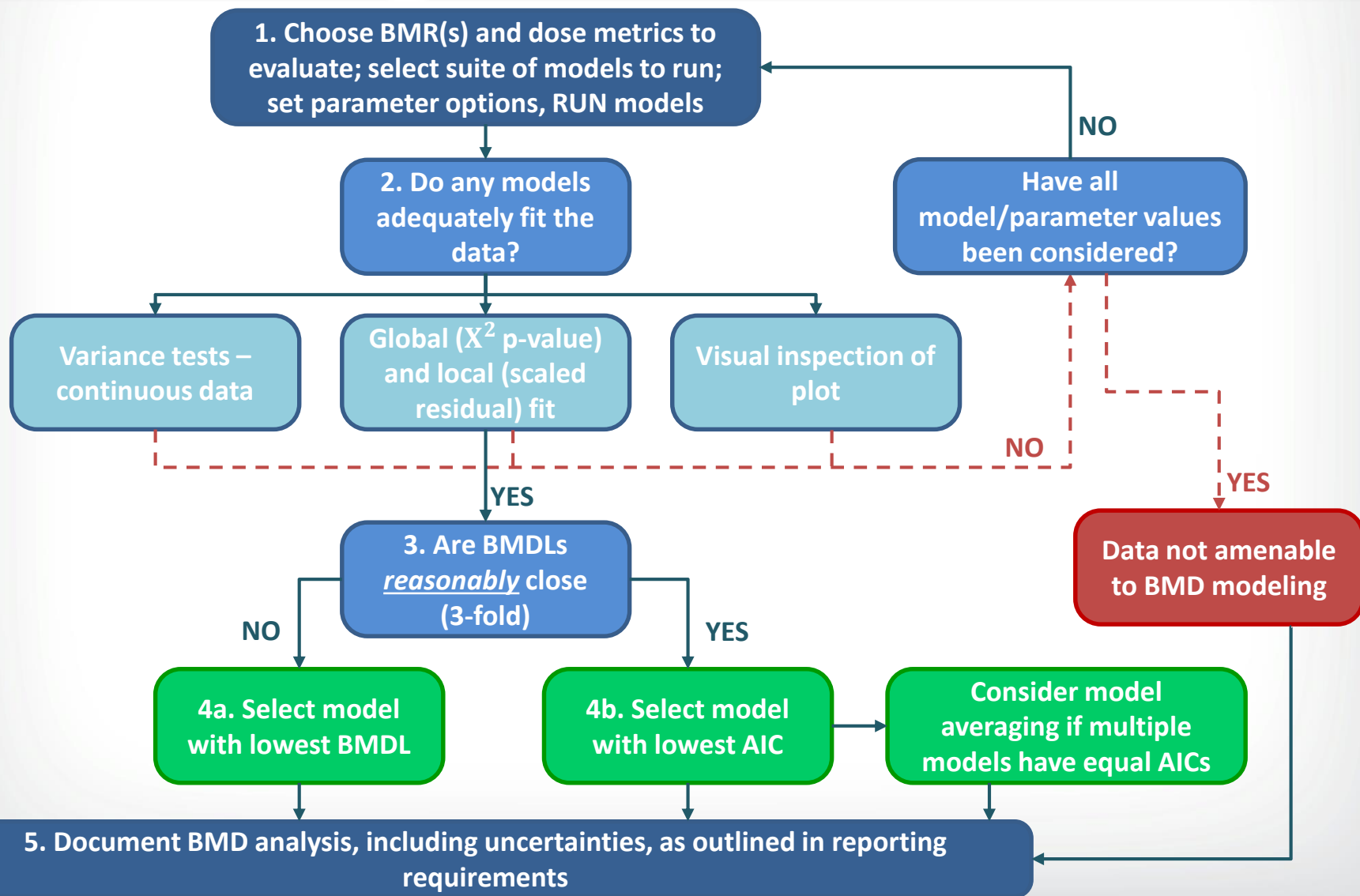
- **EPA's Benchmark Dose Software (BMDS, current version 3.1.2) is a freely-available, open-source dose-response modeling product primarily for analysis of toxicological data**
- **BMDS 3.1.2:**
 - **Dichotomous data** – data measured as binary responses (non-cancer and cancer)
 - Suite of traditional Maximum Likelihood (MLE) and Bayesian models
 - Bayesian model averaging
 - Multi-tumor analysis (MS-Combo model)
 - **Continuous data** – data measured on some continuous scale
 - Suite of traditional MLE models
 - **Nested dichotomous data** – binary, clustered responses
- **BMDS 3.2:**
 - Will implement Bayesian models and model averaging for continuous data
- **BMDS 2.7 (Archive):**
 - Repeated response data – continuous data measured at multiple time points



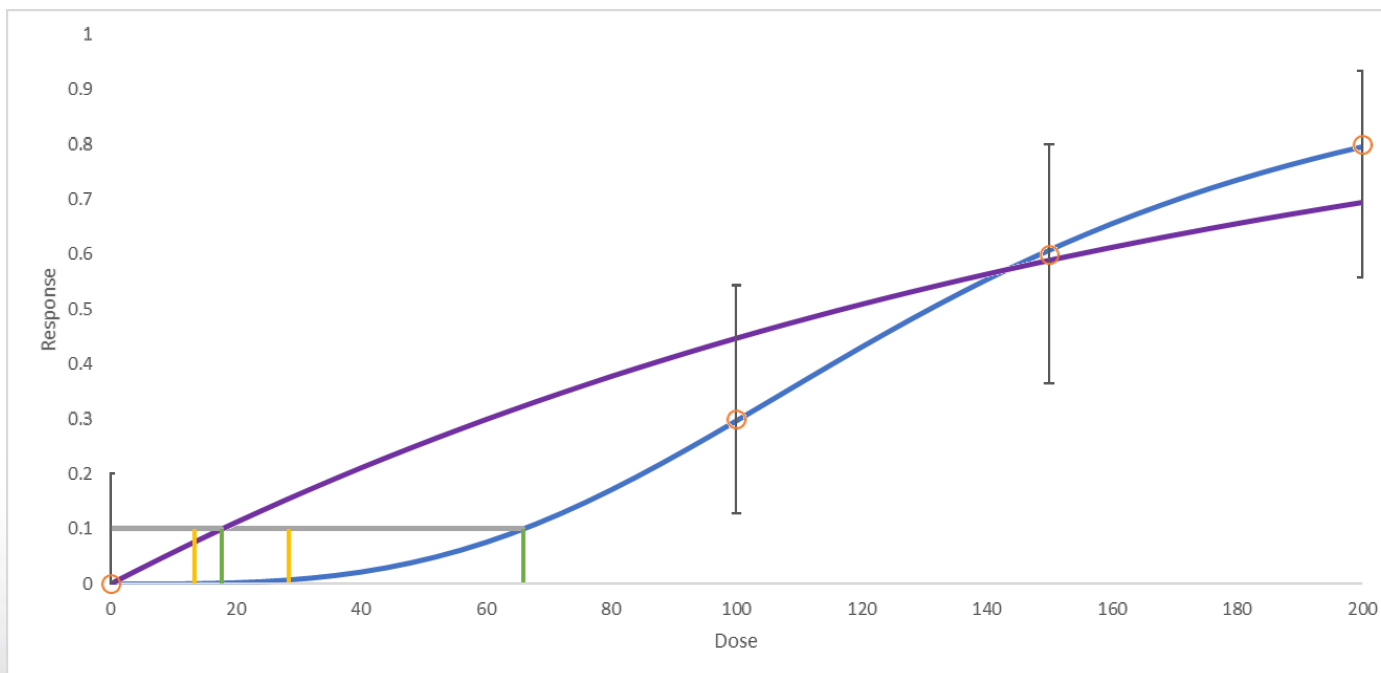
Types of Data

Data	Description	Examples	Model Inputs
Dichotomous	<ul style="list-style-type: none">• Response is measured as on/off or true/false• BMDS can only model positive dose-response trends, where incidence increases with dose	<ul style="list-style-type: none">• Tissue histopathology (non-cancer)• Tumor incidence	<ul style="list-style-type: none">• Dose• Number of Subjects• Incidence OR Percent Affected
Continuous	<ul style="list-style-type: none">• Response is measured on a continuous spectrum• Response is a numerical value with a measure of variability (i.e., standard error or standard deviation)• Response can either increase or decrease with dose	<ul style="list-style-type: none">• Body weight• Organ weight• Enzyme Activity	<ul style="list-style-type: none">• Dose• Number of Subjects• Mean response (per dose group) OR individual animal responses• A measure of variability in response (standard deviation or standard error; standard deviation automatically calculated when entering individual responses)

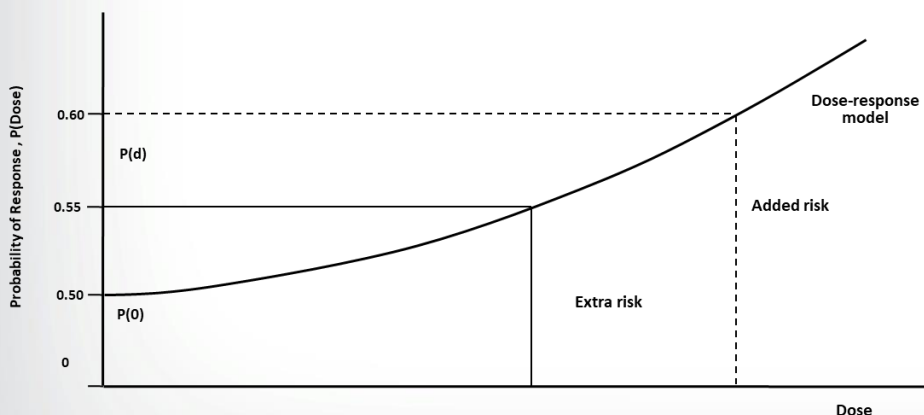
BMD Analysis – Five Steps



- **BMR** should be near the low end of the observable range of increased risks in a bioassay
- **BMRs** that are too low can impart high model dependence
- **Model dependence** can affect **BMDL** estimation such that **BMDLs** are based on model behavior and not the observed data



- For dichotomous data, BMRs are expressed as:
 - Added risk** – $AR(d) = P(d) - P(0)$
 - Extra risk** – $ER(d) = [P(d) - P(0)]/[1 - P(0)]$
- Extra risk is recommended by the IRIS Program, and is used in IRIS risk assessments.**



10% Added Risk

$0.10 = P(d) - P(0)$; if $P(0) = .50$

$$P(d) = 0.10 + P(0) = 0.10 + 0.50 = \mathbf{0.60}$$

10% Extra Risk

$0.10 = [P(d) - P(0)]/[1 - P(0)]$; if $P(0) = .50$

$$P(d) = 0.10 \times [1 - P(0)] + P(0) = (0.10 \times 0.50) + 0.50 = \mathbf{0.55}$$

The dose will be lower for a 10% Extra risk than for a 10% Added risk if $P(0) > 0$



Dichotomous BMR Selection

- **An **extra risk** of 10% is recommended as a standard (not default) reporting level for dichotomous data.**
 - Customarily used because it is at or near the limit of sensitivity in most cancer bioassays and in non-cancer bioassays of comparable size
- **In some situations, use of different BMRs is supported**
 - Biological considerations sometimes support different BMRs (5% for frank effects, >10% for precursor effects)
 - When a study has greater than usual sensitivity, a lower BMR can be used (5% for developmental studies)
 - Results for a 10% BMR should always be shown for comparison when using different BMRs.

Continuous BMR Types

BMR Type	BMR Calculation
Standard Deviation:	$\mu_0 \pm (BMRF \times SD_0)$
Relative Deviation:	$\mu_0 \pm (BMRF \times \mu_0)$
Hybrid Approach:	<p>Increasing dose-response: $\frac{\Pr(X > X_0 D) - \Pr(X > X_0 0)}{1 - \Pr(X > X_0 0)}$</p> <p>Decreasing dose-response: $\frac{\Pr(X < X_0 D) - \Pr(X < X_0 0)}{1 - \Pr(X < X_0 0)}$</p>

Where:

μ_0	=	Modeled mean response at control dose
SD_0	=	Modeled standard deviation at control dose
$BMRF$	=	BMR factor (user input used to define BMR)
$\Pr(X < X_0 0)$ or $\Pr(X > X_0 0)$	=	Background probability that defines adverse response



Using Relative Deviation as the BMR Type

- Preferred approach is to select a **BMR** that corresponds to a level of change that represents a **minimal biologically significant response** (i.e., 10% decrease in body weight, based on the **model-estimated** control mean)
- When using **RD** as the basis for the **BMR**, the user must check that the model-estimated control mean approximates the observed control means; if not, the **BMD** could be misspecified
- Consequence of using **BMRs** based on relative deviation is that the extra risk corresponding to the change can be quite high (50% by default)



Using SD as the BMD for Continuous Data

- Often, information on what response is a **minimal biologically significant response** is lacking
- 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 situations, use of different BMRs is supported**
 - For more severe effects, a BMR of 0.5 SD can be used
 - Results for a 1 SD BMR should always be shown for comparison when using different BMRs.

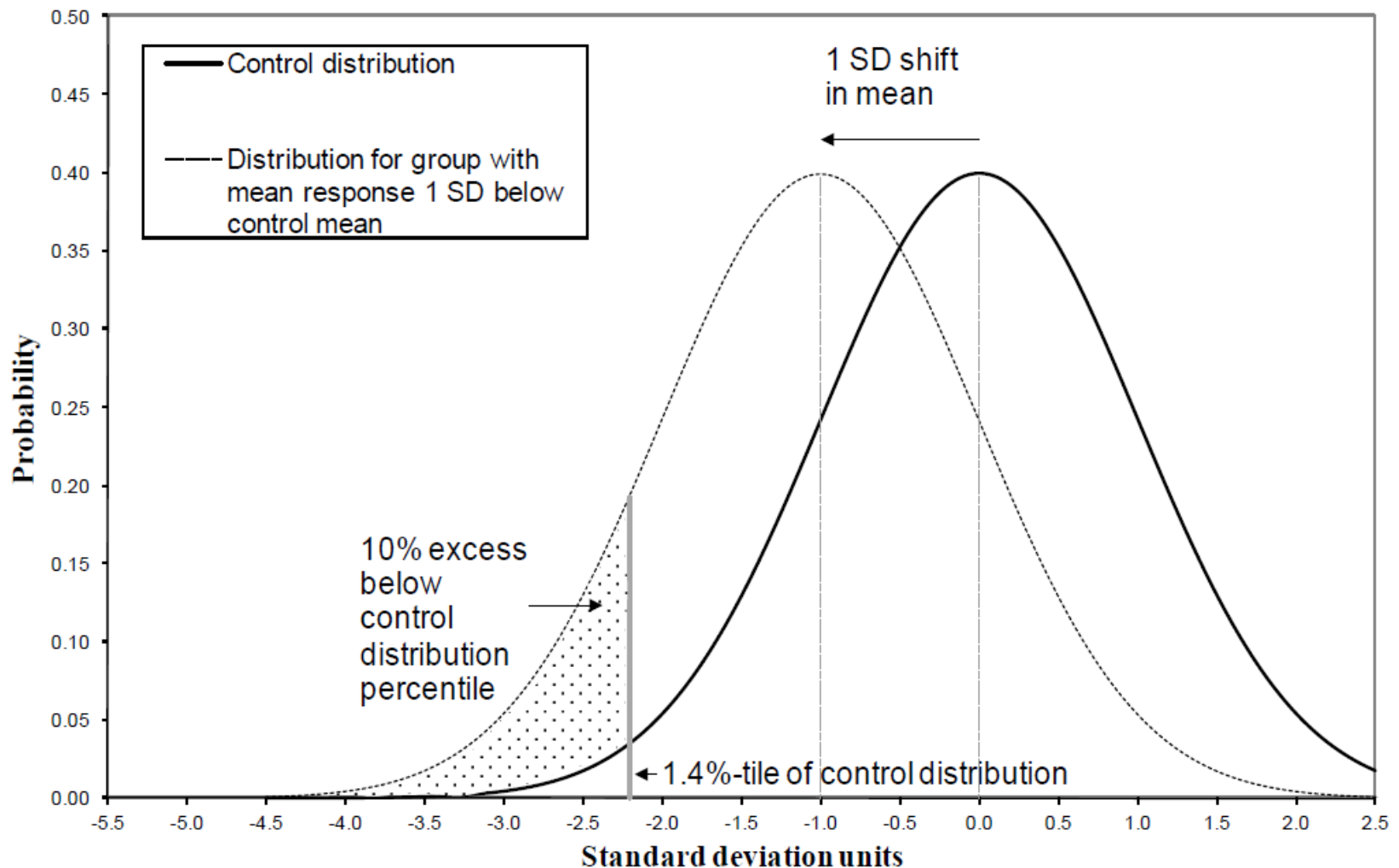


Why Use SD as the BMR for Continuous Data?

- **For a continuous endpoint in a normally distributed population, if**
 - 1.4% of the animals in the control group are assumed to have an “abnormal response,” a change in the mean response by one standard deviation will result in 10% of the animals reaching the abnormal response level (Crump, 1995)
 - This response in 10% of the animals is comparable to the 10% BMR used in dichotomous data modeling
- **NOTE: This assumes a simple shift in a normal distribution. Some toxicity responses may not behave this way**



Why Use SD as the BMR for Continuous Data?

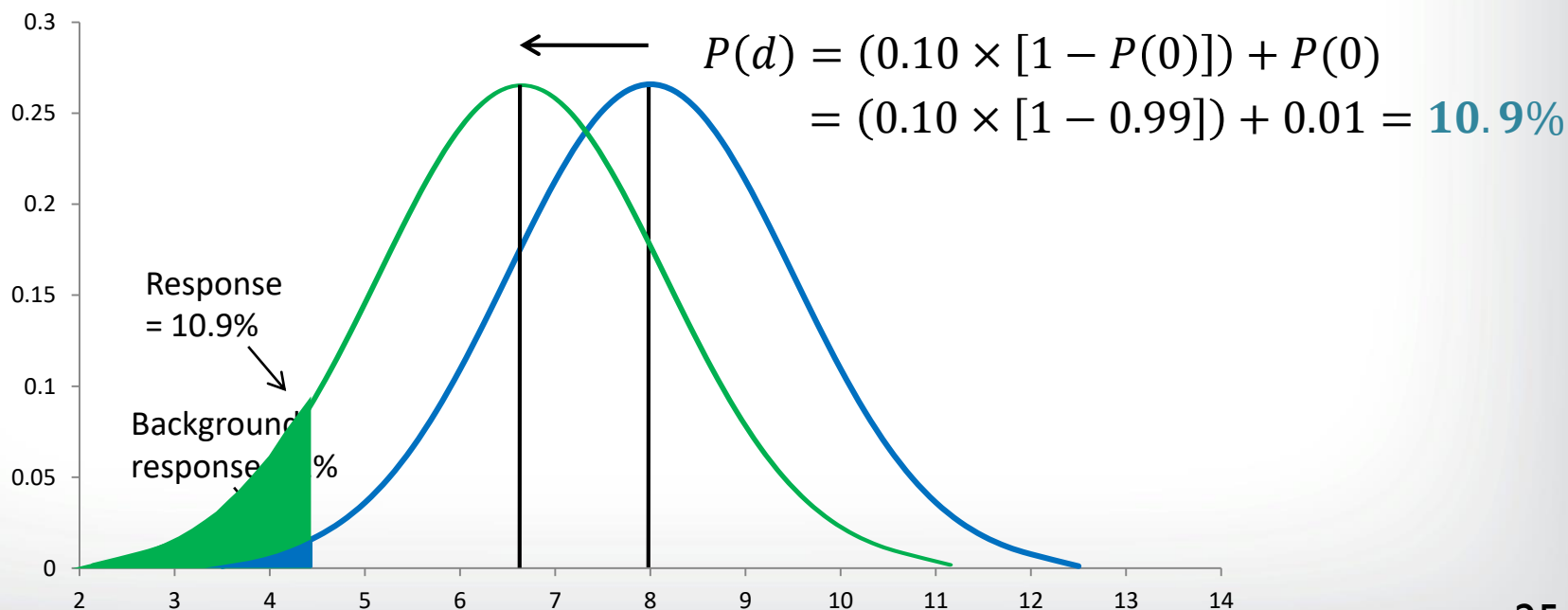




The Hybrid Approach for Continuous BMRs

- **The “hybrid approach” is an alternative method for selecting a BMR in order to calculate a BMD for continuous data**
- **Using the hybrid approach, risk is expressed in the same manner as with dichotomous models – as added or extra risk.**
- **Two parameters must be selected by the user:**
 - The benchmark response (BMR) – expressed as either added or extra risk (e.g., 10% extra risk)
 - The background rate (i.e., probability) of an adverse response in the control group

- Consider at **BMR = 10%** and a background rate = **1%**
- Model calculates the cut-off values in the control group distribution that correspond to the background rate
- Model calculates the dose that corresponds to a shift in the mean that results in **10.9%** of the animals falling beyond the cut-off values



Selection of a Specific Model

Biological Interpretation	<p>Examples:</p> <ul style="list-style-type: none">• Dichotomous:<ul style="list-style-type: none">• Saturable processes demonstrating Michaelis-Menten kinetics (Dichotomous Hill model)• Two-stage clonal expansion model (cancer endpoints)• Continuous:<ul style="list-style-type: none">• Can use the Hill or Exponential models for receptor-mediated responses
Policy Decision	<ul style="list-style-type: none">• U.S. EPA's IRIS program uses the multistage model for cancer data (i.e., dichotomous data)<ul style="list-style-type: none">• sufficiently flexible to fit most cancer bioassay data• provides consistency across cancer assessments• U.S. EPA's OPP group uses the Exponential models for modeling acetylcholinesterase inhibition data
Otherwise	<p>However, 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</p>

Dichotomous Models

Model name	Functional form	# of Parameters ^a	Low Dose Linearity	Model fits
Multistage	$\gamma + (1 - \gamma) \left[1 - \exp \left\{ - \sum_{j=1}^k \beta_j X^j \right\} \right]$	1+k	Yes, if $\beta_1 > 0$ No, if $\beta_1 = 0$	All purpose
Logistic	$\frac{1}{1 + \exp\{-(\alpha + \beta X)\}}$	2	Yes	Simple; no background
Probit	$\Phi(\alpha + \beta X)$	2	Yes	Simple; no background
Log-logistic	$\frac{\gamma + (1 - \gamma)}{1 + \exp\{-[\alpha + \beta \ln(X)]\}}$	3	No	All purpose; S-shape with plateau at 100%
Log-probit	$\gamma + (1 - \gamma) \Phi\{\alpha + \beta \ln(X)\}$	3	No	All purpose; plateau S-shape with plateau at 100%
Gamma	$\gamma + (1 - \gamma) \left[\int_0^{\beta x} t^{\alpha-1} e^{-t} dt \right] / \Gamma(\alpha)$	3	No	All purpose
Weibull	$\gamma + (1 - \gamma)[1 - \exp\{-\beta X^\alpha\}]$	3	No	"Hockey stick" shape
Dichotomous Hill	$v \times g + \frac{(v - v \times g)}{1 + \exp\{-a - b \times \ln(X)\}}$	4	Yes	Symmetrical, S-shape with plateau

^a Background parameter = γ . Background for hill model = $v \times g$

Model Name	Functional Form	# of Parameters	Model Fits
Polynomial ^a	$\beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_n X^n$	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	$a \times \exp\{\pm 1 \times b \times X\}$	2	All purpose (Models 2 & 3)
Model 3	$a \times \exp\{\pm 1 \times (b \times X)^d\}$	3	Symmetrical and asymmetrical
Model 4	$a \times [c - (c - 1) \times \exp\{\pm 1 \times b \times X\}]$	3	S-shape with plateau (Models 4 & 5)
Model 5	$a \times [c - (c - 1) \times \exp\{\pm 1 \times (b \times X)^d\}]$	4	

^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 Parameters in Dichotomous MLE Models

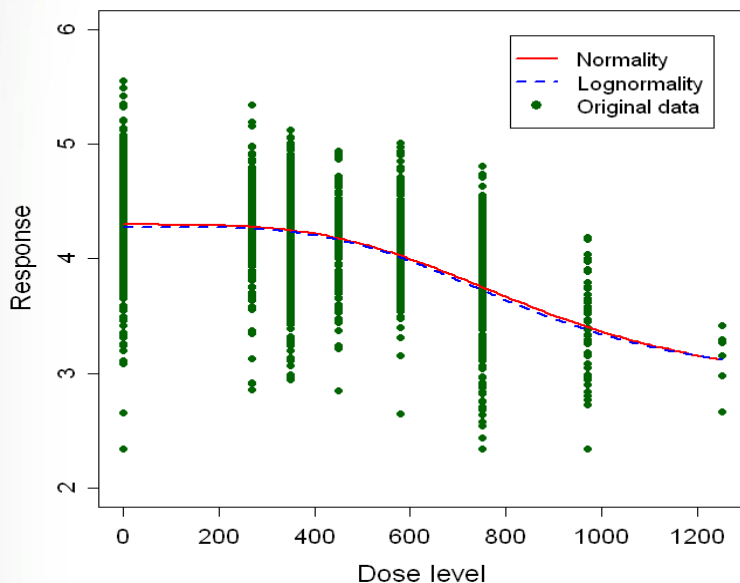
- **Dichotomous MLE models are conceptually restricted so that probabilities are positive numbers no greater than one**
- **By default, BMDS models are restricted to prevent biologically implausible curve shapes**
 - For instance; power parameters can be restricted to be ≥ 1 and < 18
 - See BMDS User Guide for details on EPA preferred default model restrictions
- **These MLE model restrictions can impact statistical calculations such as the goodness-of-fit p-value and AIC**
 - Currently, a parameter estimate that “hits a bound” impacts a model’s degrees of freedom (DF) (in BMDS, DF is increased by 1 for p-value calculation)
 - When a parameter hits a bound, that parameter is not counted towards the AIC penalization (EPA’s Statistical Working Group may modify this approach in the future)

- **Data can be assumed to be normally or lognormally distributed for continuous data**
 - This reflects the distribution of the data per se, not how the modeling is done
 - Many biological parameters are lognormally distributed; a lognormal distribution is also useful to consider whenever responses are constrained to be positive
 - When using summary data (observed means and SD), modeling with a log-normal distribution gives an approximate maximum likelihood estimate
 - The SD is homogenous on a log-scale when within dose-group variance is proportional to the mean response
- **An extra parameter is needed to model the within dose-group variance if normality is assumed**
- **Sometimes, the extra parameter can have significant impact on the BMD estimation if the “Hybrid” approach is used (Shao et al., 2013)**

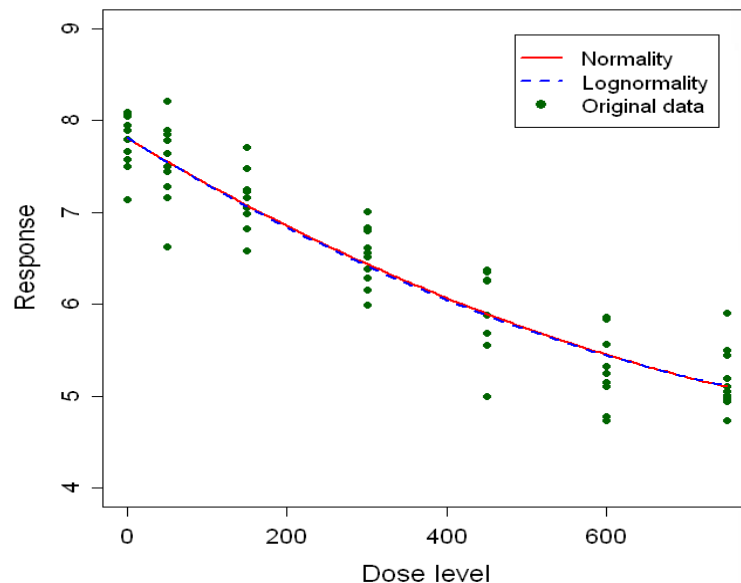


Continuous Model Distributions

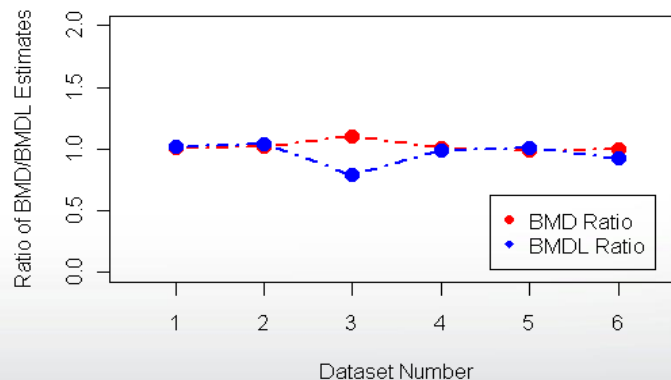
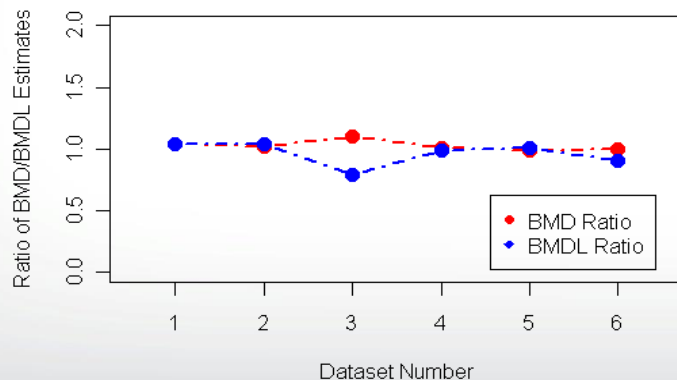
Hill Model



Quadratic Model



Ratios of BMD and BMDL Estimates from Quadratic Model using 5%(left) and 10%(right) Relative Deviation



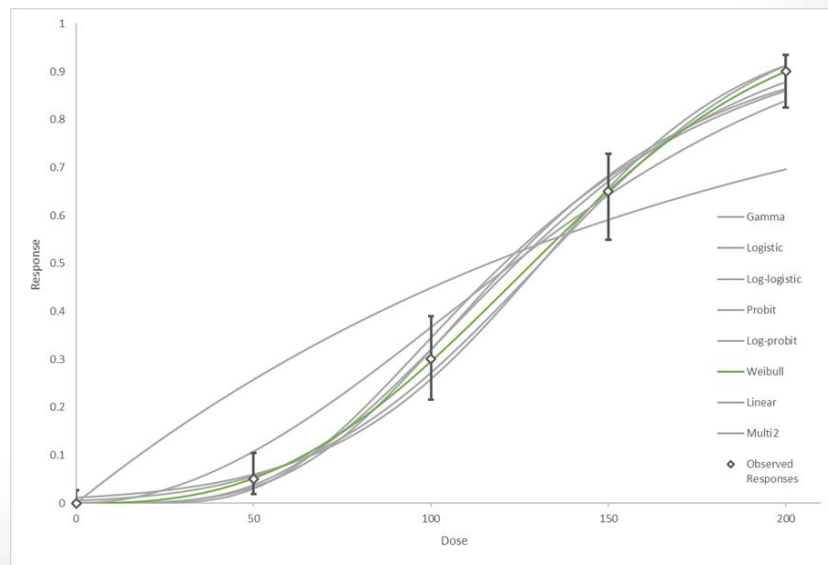
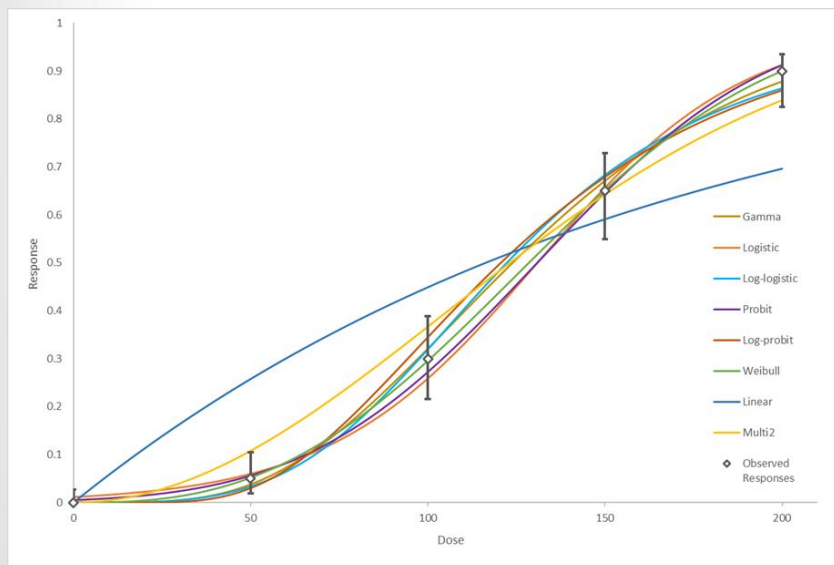


Does the Model Fit the Data?

- **Tests of interest (response/variance modeling) (*continuous MLE models only*)**
 - Test 1 cut-off: $p < 0.10$
 - Tests 2 and 3 cut-off: $p > 0.05$
- **Global measurement: goodness-of-fit p value ($p > 0.1$) (*MLE models only*)**
- **Local measurement: Scaled residuals (absolute value < 2.0) (*MLE models only*)**
- **Visual inspection of model fitting.**

- **Often, more than one model or modeling options will result in an acceptable fit to the data.**
- **When comparing models from different families, Akaike’s Information Criterion (AIC) is a commonly used method to identify the best fitting model (the lower the AIC, the better)**
 - $AIC = -2 \times LL + 2 \times p$
 - LL = log-likelihood at the maximum likelihood estimates for parameters
 - p = number of model degrees of freedom (dependent on total number of model parameters, number of model parameters that hit a bound, and the number of dose groups in your dataset)
 - Only the DIFFERENCE in AIC is important, not actual value
- **Consider using the lowest BMDL if BMDL estimates from acceptable models are not sufficiently close, indicating model dependence**
 - What is “sufficiently close” can vary based on the needs of the assessment, but generally should not be more than 3-fold.

- **When fitting multiple models to a single dataset, many models can (*and often will*) statistically fit the data well**
 - So, is there a compelling reason (toxicology, MOA, etc.) to pick one model over any other?
 - Or (most commonly) is the model selected based on pure statistical fit?
 - This is *model uncertainty*



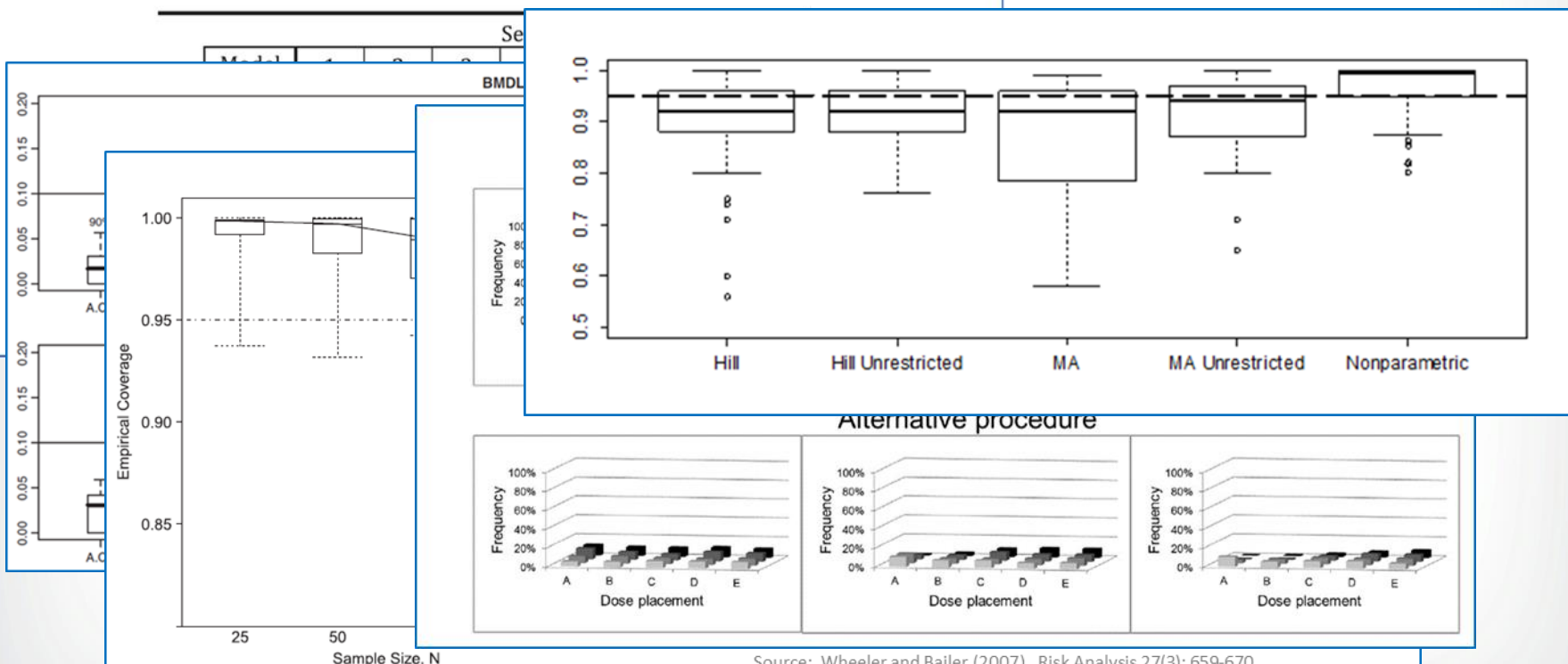
- **Multiple approaches have been developed for addressing and/or characterizing model uncertainty**
 - Flexible parametric models – some research has indicated that some models (Exponential 5) are flexible enough to fit the majority of dose-response shapes observed in the literature
 - Semi- or non-parametric models – completely data-driven models that are hyper-flexible
 - **Model averaging** – methods by which the results of a suite of individual models are averaged together to give one estimate of the BMD and BMDL



Why Move Away from Single Model Selection?

- Research convincingly shows that single model selection practices are often sub-optimal compared to model averaging

Table 4. AIC-based model selection percentages over all configurations (A–F) in Table 2 for per-dose sample size of $N=50$



Source: Piegorsch et al. (2013). *Environmetrics* 24(3): 143-157

Source: Wheeler and Bailor (2007). *Risk Analysis* 27(3): 659-670



BMDS Bayesian Model Averaging

- **Unique aspects of EPA/NIOSH model averaging approach:**
 - Informed priors
 - Based on knowledge of how chronic studies are designed and where the BMD_{10} estimates are most likely to be relative to a study's maximum dose
 - Disentangle issues related to models that “degenerate” to other models (Weibull, Gamma, etc.)
 - Prevent over-fitting of individual models
 - Provides a single standard set of priors in the “non-research” version of BMDS (i.e., Excel version) that gives reasonable, health-protective, consistent, and reproducible results
 - Laplace approximation of posterior density
 - Minimal loss of accuracy or reliability
 - Substantial increase in speed (~10-fold faster than MCMC approaches implemented in other platforms)
 - Increases in speed are critically important for batch analyses of many datasets



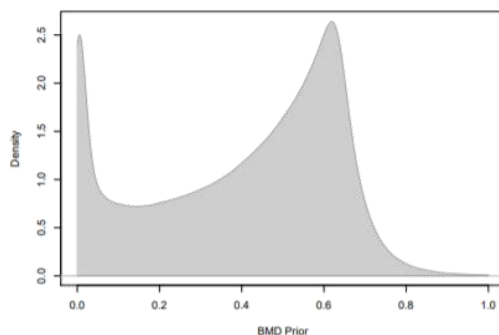
Priors on a Dose-Response Curve

- **For setting priors over a dose-response curve, there are many options**
- **Two that have been published on in the literature are:**
 - **Flat Priors:** In the search for an objective prior, the selection of a prior that is uniform over some range can be used (Shao and Shapiro, 2018)
 - **Focused Prior:** focus on building a prior on a value of inferential importance. In this case, the value of interest is the benchmark dose (Fang et al., 2017, BMDS 3.2, 2019)
- **In judging the behavior of these prior options, the focus is on the maximum tolerated dose (MTD) as there is a large literature about the correlation between MTD and the point of departure (NRC, 1993)**

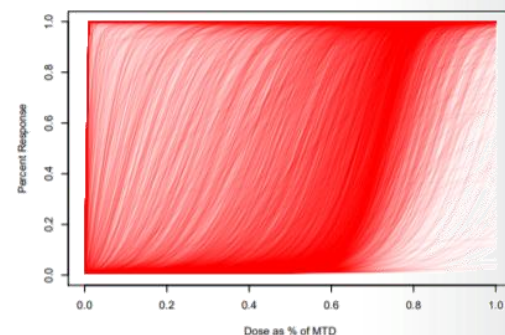
- **Place uniform priors over standard dose-response model parameters**
- **Benefits:**
 - This is most like previous BMD analyses (i.e., maximum likelihood estimation with bounds set on model parameters)
 - If the true parameter is in the bounds (of the prior), the true parameter value will be obtained as $n \rightarrow \infty$
 - Analysis is not biased in most cases
- **Issues:**
 - Does not necessarily generate dose-response curves that look like real data
 - Are not created based on the parameter of interest; i.e., the BMD
 - Can bias analyses in “edge” cases

- For the Weibull model, a flat prior has behavior that may not be intuitive in terms of generating data and the BMD (in terms of the MTD)
- One further issue with flat priors is the selection of the bounds: the flatter (or more uniform) the prior, the more biased towards the MTD the BMD is

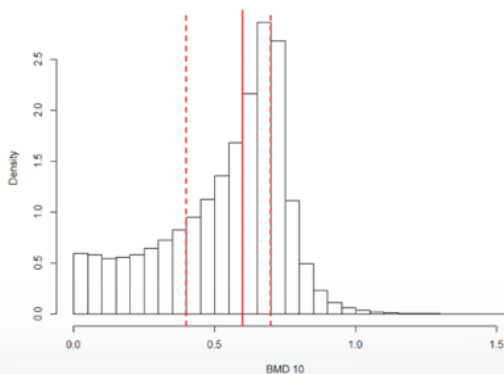
Prior on the BMD



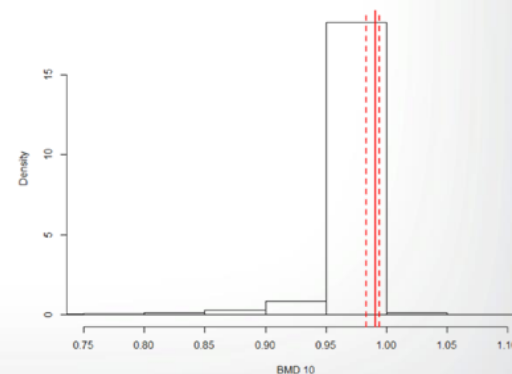
Prior on the DR curves.



Weibull Flat prior $\alpha \sim \text{Uniform}(1,15)$



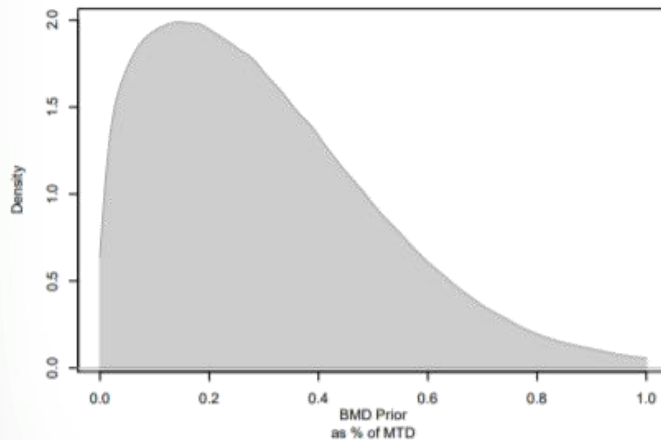
"Super Uninformative" Prior $\alpha \sim \text{Uniform}(0.5,1000)$



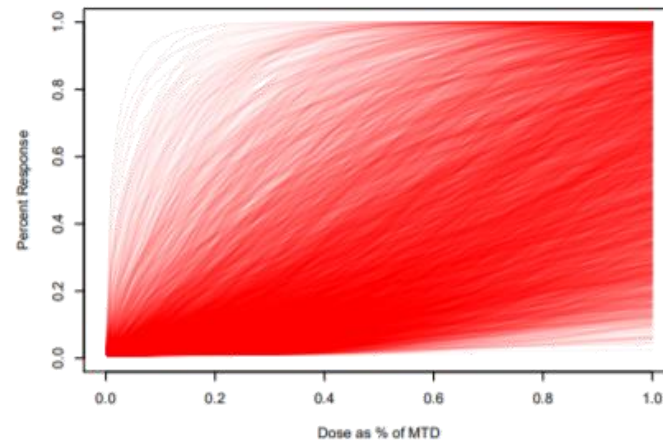
- **Instead of looking at priors over all model parameters, or specific parameters, place a reasonable prior over the value of ultimate interest, the BMD**
- **Benefits:**
 - All models are wrong, so the parameters are abstract entities. We are ultimately interested in the value of the BMD
 - In terms of MTD and dose-response study design, the value of the BMD can be expressed as a percentage of the MTD.
- **Issues:**
 - Can be perceived as subjective in terms of what is “right”
 - Significant prior impact in low data cases
 - Might change based upon target quantity (i.e., may be different for BMR = 10% vs. BMR = 1%)

- Here, assuming a prior on the **BMD** such that the majority is between 0 and 0.5 of the **MTD**
- Result is similar to Informative priors to the null, but the variability in the possible curves shapes is reduced

Prior on the BMD



Prior on the DR curves.



- In **BMDs 3.1.2** priors for dichotomous models are set such that the **BMD** is expected to fall within $\sim 0.2-0.5$ of the **MTD**



Parameter Constraints vs. Priors

- In MLE versions of models, hard constraints are placed on parameters to prevent certain curve shapes
- For example, the Weibull model:

$$Weibull = \gamma + (1 - \gamma)(1 - \exp[-\beta d^\alpha])$$

- Constraint often put on α parameter (≥ 1) to prevent supralinear linear curves
- When parameters are estimated on boundary, statistical inferences are impacted
- Bayesian models replace hard-constraints with parameter priors that *place low prior probabilities on certain parameter values*
 - Prior for α parameter is: $\log(\alpha) \sim \text{Normal}(\log(2), 0.18)$
 - This corresponds to a very low probability that the value of the α parameter is < 1
 - So, parameter priors allow certain parameter values, but *conclusive data is required for parameter to take those values*

- **Analytical means are necessary to fit dose-response models to the observed data and estimate the posterior distribution**
- **MCMC (Markov Chain Monte Carlo) – Gold standard: this method uses sampling from the posterior distribution using a method that converges to that distribution**
 - Will give the posterior distribution
 - Can never know if converges to the target distribution
 - Can take time and is more complicated than finding the maximum
- **Maximum a-posteriori – find the maximum of the posterior distribution and use a normal like approximation**
 - Don't know the size of the sample that adequately approximates the posterior
 - Very fast computationally compared to MCMC
 - Accurate for the right-sized posterior

- **Prior model averaging methods used AIC or BIC (Bayesian Information Criterion) as weights in the averaging**
- **BMDS instead uses the Laplace approximation to the marginal density of the data and weights are calculated as**

$$\pi_k(M_k|D) = \frac{f(M_k)I_k}{\sum_{i=1}^n f(M_i)I_i}$$

- **The model-averaged BMD point estimate is the weighted average of the MAP estimates from individual models**
- **BMDL and BMDU values are estimated similar to the profile likelihood approach except that the posterior density is profiled**



BMDS Bayesian Model Averaging

- **EPA/NIOSH BMA** approach was extensively tested against 1) **MCMC Bayesian MA** approach with uninformative priors; 2) **BMDS** using 2012 model selection criteria; and 3) flexible non-parametric model
- **34 separate “true-dose” curves** used to test approaches

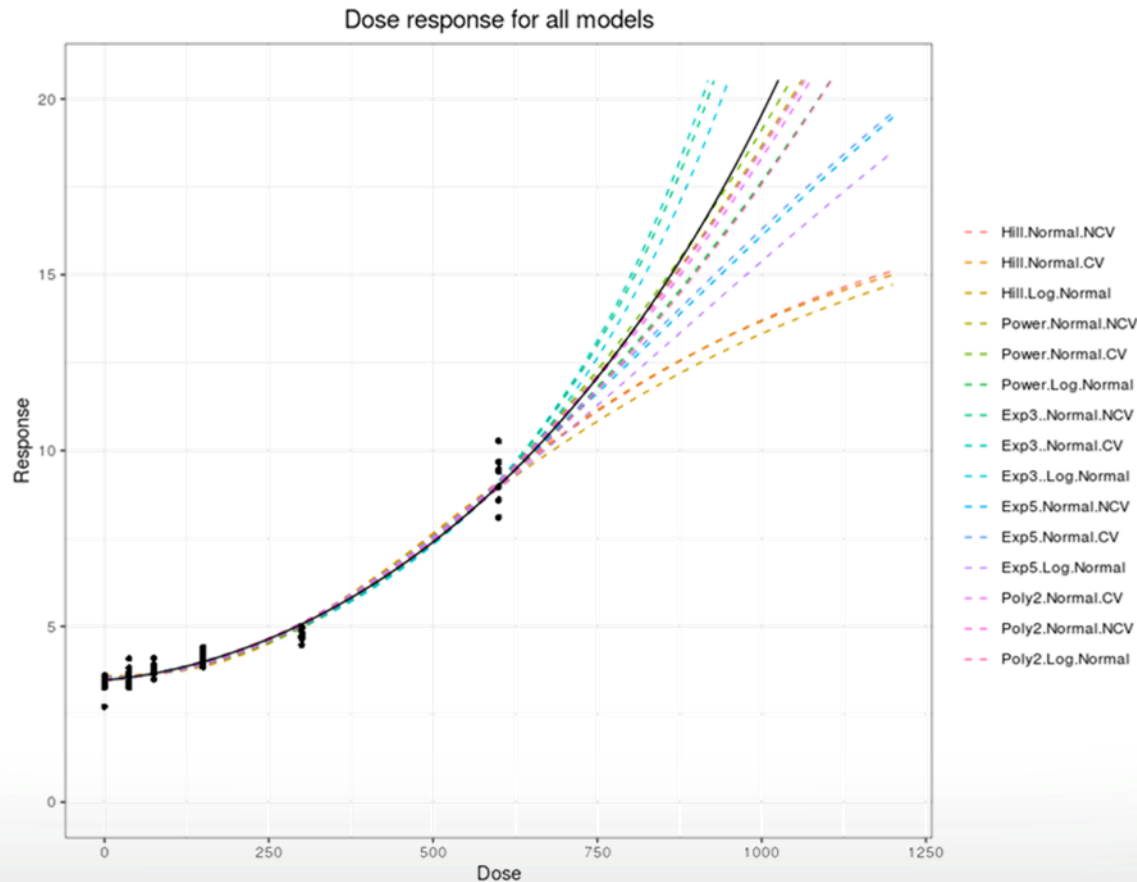
Percentage of Times BMDL Coverage is >90% than True BMD Value

True BMD	BMA	BMDS	NP	MCMC
All templates	70.6%	41.2%	76.5%	47.1%
True BMD < 0.2x max dose	63.2%	26.3%	57.9%	36.8%
True BMD > 0.2x max dose	80%	60%	100%	60%
True BMD < 0.1x max dose	60%	30%	20%	40%
True BMD > 0.1x max dose	75%	45.8%	100%	50%

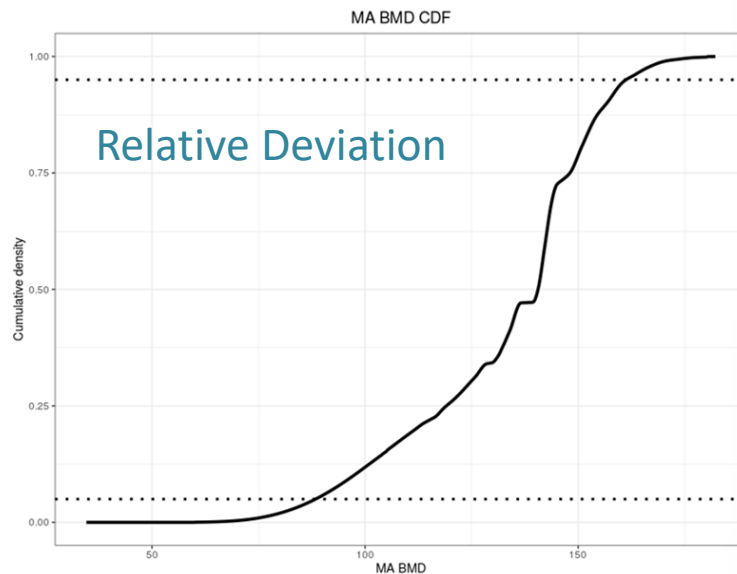
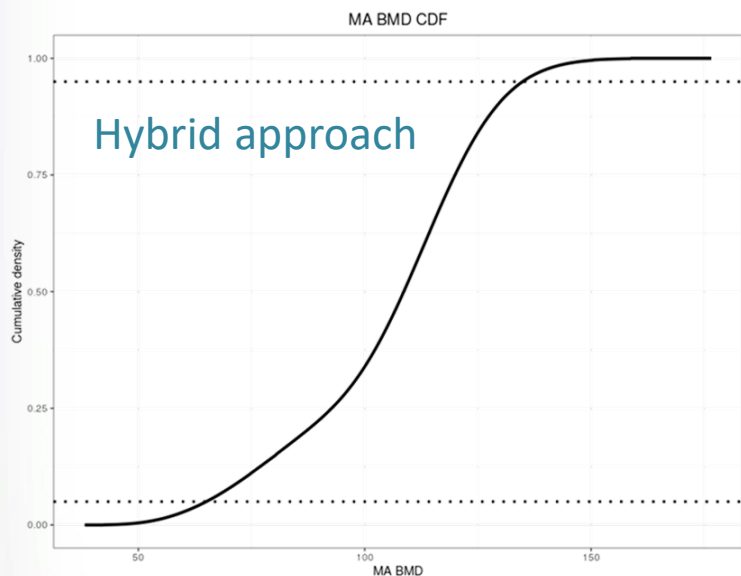
- **Continuous model averaging poses a different problem from dichotomous model averaging:**
 - Slob and Setzer (2014) showed that using the Exponential 5 or Hill model was usually adequate for fitting a wide array of dose-response shapes
 - Shao and Gift (2013) showed that the BMDs estimated using a BMR based on relative deviation is essentially the same when using either the normal or log-normal distribution
- **So, why do we even need model averaging for continuous endpoints?**

- There are cases where the distribution does actually make an impact in the modeling results:
 - Using the standard deviation definition of the benchmark response will result in different benchmark doses based upon the assumed distribution
 - The same is true for the hybrid approach
- Using model averaging approaches, there is no reason one can not average over models and distributions
 - The assumed distribution is technically part of the model too.

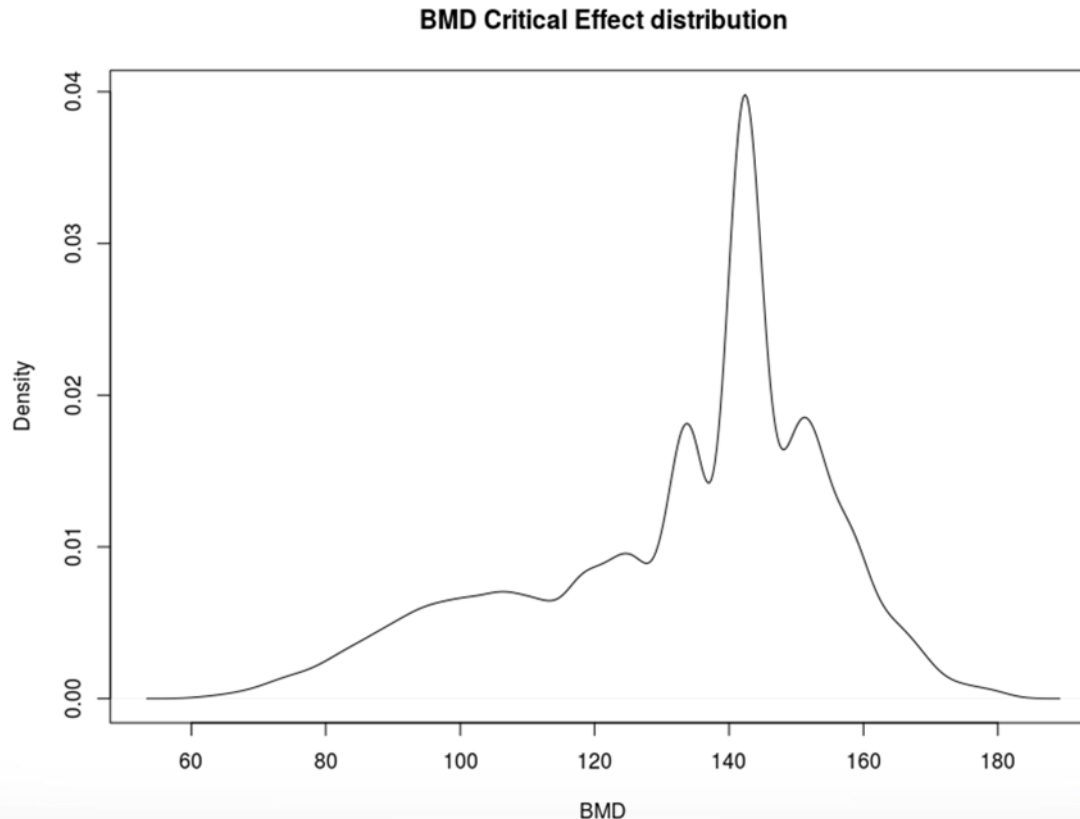
- Looking at the mean response, everything seems similar to dichotomous model averaging



- But things change when you look at the **CDF**

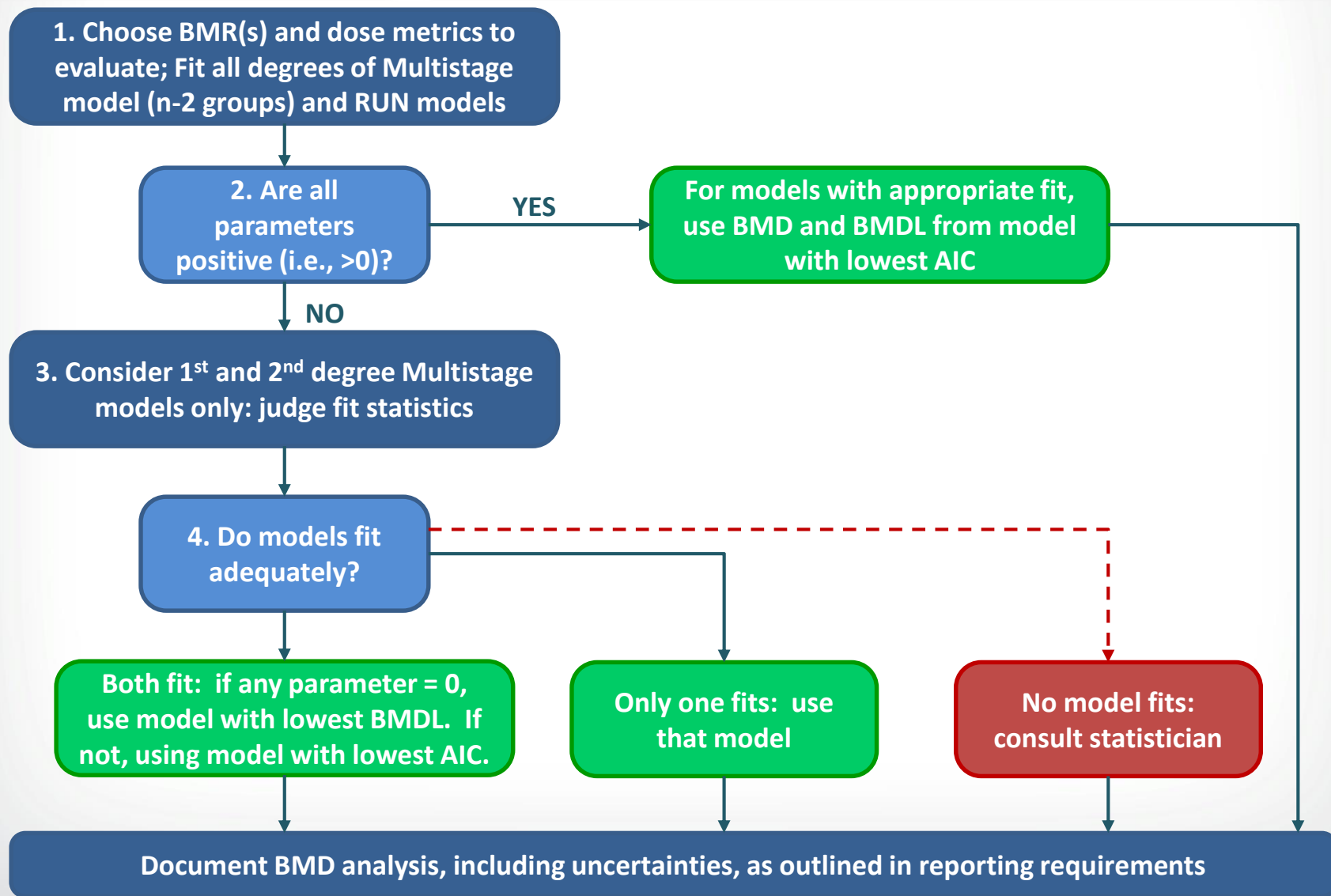


- The strange **Relative Deviation CDF** is due to the multiple modes from the model average



- **Benefits:**
 - Up to 24 model-distribution-variance combinations included in averaging suite: up to eight models (Exp2, Exp3, Exp 4, Exp5, Hill, Power, Poly2, Linear) × three different distribution/variance combinations (normal – constant variance, normal – non-constant variance, log-normal-constant variance)
 - Model averaging provides a better picture of uncertainty than using one flexible parametric model
- **The ultimate selection of priors for continuous models is still being researched for BMDS.**
 - Priors are generally diffuse
 - Are designed to prevent drastic (i.e., non-biologically plausible) on/off responses

Description	<ul style="list-style-type: none"> • Response is measured as on/off or true/false • You either have it or you don't • BMDS can only model positive dose-response trends, where incidence increases with dose
Example Endpoints	<ul style="list-style-type: none"> • Cancer: Tumor incidence
Model Inputs	<ul style="list-style-type: none"> • Dose • Number of Subjects • Incidence or Percent Affected
Model	<ul style="list-style-type: none"> • Multistage Cancer: <ul style="list-style-type: none"> • β coefficients always restricted to be positive • Cancer slope factor calculated • Linear extrapolation shown on plot • Form: $\gamma + (1 - \gamma) \left[1 - \exp \left\{ - \sum_{j=1}^k \beta_j X^j \right\} \right]$



- **Often, an individual cancer bioassay will report dose-related increases in multiple, independent tumor types**
 - Basing unit risk estimates on only one tumor type may underestimate the carcinogenic potential of a chemical that is observed to induce neoplasia at multiple sites in a bioassay (NRC, 1994)
- **MS-Combo model allows users to calculate the BMD and BMDL for *any combination* of tumors observed in a single bioassay.**
- **The major assumption of the MS-Combo model is that different tumor types are *independent* of one another**
 - Independence can be determined based on statistical or biological considerations
- **Individual tumor types must first be modeled with the multistage model to determine which degree model best fits the data**
 - This allows individual tumors to be fit with models that best characterize their specific dose-response shapes



Poly-3 Survival Adjustment

- **The poly-3 survival adjustment is a method to calculate survival-adjusted lifetime tumor rates by fractionally weighting the number of exposed animals (i.e., sample size)**
 - “Poly-3” refers specifically to using a 3rd order polynomial to describe the tumor incidence function in time
 - Other polynomials can be used, but estimating the correct polynomial can be difficult
 - Failure to adjust for differential mortality can bias modeling results
- **For an individual dose group (*i*), the poly-3 survival adjusted sample size is:**

$$n_i^* = \sum_{j=1}^{n_i} w_{ij}$$

- Where, $w_{ij} = 1$ if the *j*th animal in the *i*th dose group had a tumor at observation (i.e., necropsy); otherwise, $w_{ij} = t_{ij}^3$, where t_{ij} is the fraction of duration of the study for which the animal survived



Developmental Toxicity Data

Data	Description	Examples	Model Inputs
Dichotomous	<ul style="list-style-type: none">• Fetal response (on/off) reported for individual exposed dams• BMDS can only model positive dose-response trends, where incidence increases with dose	<ul style="list-style-type: none">• Malformations• Fetal death	<ul style="list-style-type: none">• Dose• Number of fetuses at risk for each litter (i.e., <i>individual litter level data required</i>)• Number of fetuses affected for each litter
Continuous	<ul style="list-style-type: none">• Response is measured on a continuous scale• Response is a numerical value with a measure of variability• Response can either increase or decrease with dose• Responses are measured for fetuses within individual litters	<ul style="list-style-type: none">• Fetal body weight• Fetal organ weight• Pup weight gain PND 7-14	<ul style="list-style-type: none">• Dose• Number of Subjects (litters or fetuses)• Mean response (mean of litter means; per dose group) OR individual fetal responses• A measure of variability in response (standard deviation or standard error; standard deviation automatically calculated when entering individual responses)



Modeling Developmental Toxicity Data

- **Must account for the litter effect; the propensity of litter-mates to respond more alike one another compared to offspring from different litters**
 - Failure to do so will underestimate the variances
 - Meaning, dose-response modeling results will be biased
- **For dichotomous data:**
 - Can use nested models in BMDS if individual dam (i.e., individual litter) data is available
 - If only summary data (i.e., dose group level) data is available, can use Rao-Scott transformation and regular dichotomous models
- **For continuous data:**
 - If individual fetal or litter data is available, correct variances are easy to calculate
 - If only summary data is available, approximate methods are used to correct variances
 - Regular continuous models used in both cases

- **Implementation of continuous Bayesian model averaging**
 - Approach will average over models, distributions, and variances
 - Addresses the uncertainty with having to assume a particular distribution *a priori*
- **Release of BMDS-HAWC**
 - Interoperable online version of BMDS
 - Fully integrated into EPA's HAWC online assessment database
- **Release of BMDS-R**
 - “Research” version of BMDS
 - Fully configurable
 - Will facilitate further development of 3rd-party BMDS products
- **Continued dose-response research**
 - Model priors
 - Unified model suite for dichotomous and continuous data
 - Nested continuous model for incorporating litter specific covariates