NTP
National Toxicology Program

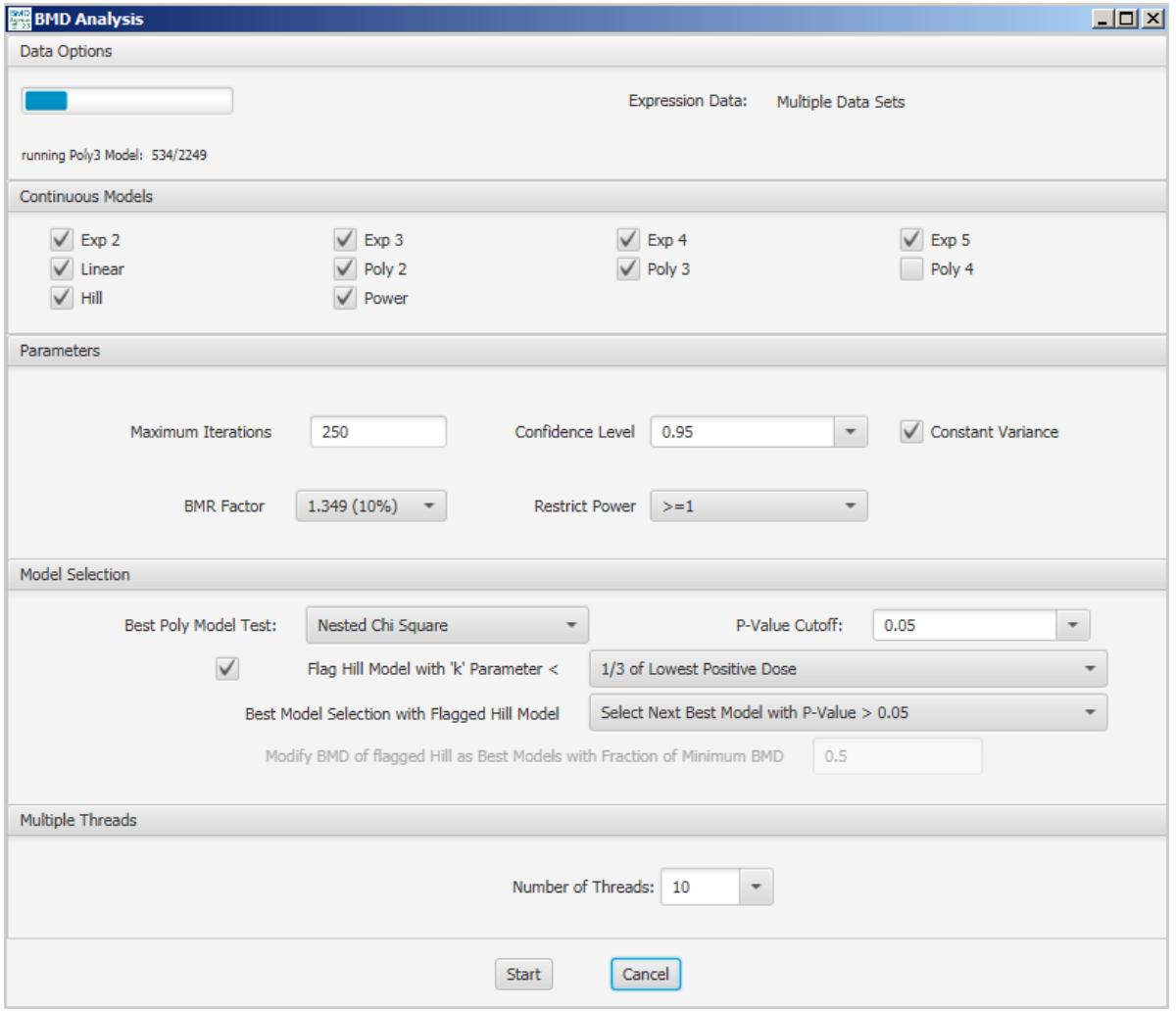
NTP's Proposed Approach to Curve Fitting and Determination of Feature Potency

Scott S. Auerbach Ph.D., DABT
Biomolecular Screening Branch
National Institute of Environmental Health Science

Expert Panel Meeting on the Peer Review of Draft NTP Approach to Genomic Dose-Response Modeling
October 24, 2017

Fitting Features to Dose-Response Curves

- Features are fit to 9 parametric continuous models
  - Derived from US EPA's BMDS software
  - Hill, Power, Linear, Poly2,3, Exp2,3,4,5
- BMR = 1.349 x SD of controls
  - Approximates 10% shift area under the normal distribution
- 2 step process for best model selection
  - Nested Chi Square – Best Poly model
  - Lowest AIC
- From the best fit model a BMD, BMD_L and BMD_U is determined

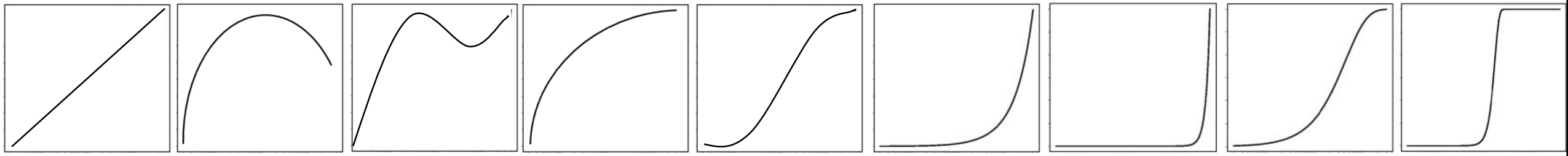


Why Parametric Models?

- Common standard for modeling toxicity data for use in risk assessment
- Simple to understand and therefore easy to translate
- Computationally affordable
- We don’t anticipate seeing significant value added by using non-parametric models because of the diversity of model types we are using

Why 9 models?

- For any given feature there is not a pre-defined response model
  - Due to the complexity associated with control of gene expression
- The model that fits the data best will provide the most accurate estimate of the BMD for that feature

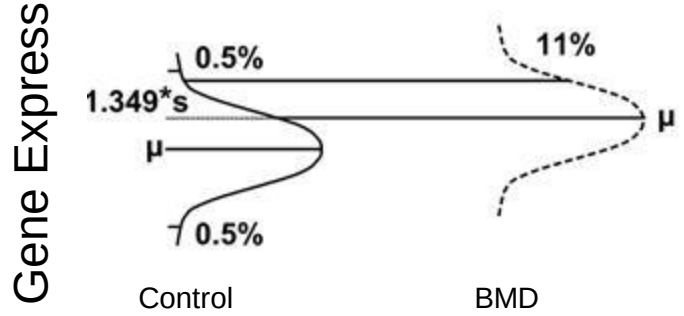


Why use a 2 Step Model Selection?

- Prefer to use a statistical method when possible such as in the case of the nested poly models, because it is a more rigorous test for comparison of fit compared to using AIC

Why use a Benchmark Response of 1.349 x SD?

- BMR = Pre-defined level change in measured feature that is used to determine the BMD
- Typically 10% response rate over control in toxicity studies
- BMR = 1.349 x standard deviation at control
- Assumptions
  - Normal distribution of intensities or counts in the control and treated groups
  - Changes in expression could occur in either tail, with a 1% chance of that occurring in the absence of exposure (0.5% in each tail)
- A shift in the mean of the treated distribution by 1.349*SD is the amount required to shift the treated distribution by 10% relative to the control distribution



Adapted from Thomas et. al., Tox Sci, 2007

Points to Consider

- Parametric models
- Fit 9 models
- A BMR of 1.349*standard deviation at control
- Two step process for model selection

- Other variables to consider
  - EC10 vs. BMR?
  - Model averaging?
  - Potential for overfitting?
  - More models and less models?
  - BMD outside the dose range, use grid search to identify value?
  - Spline-based metaregression?